Identifying How Successful Vocabulary-Learning Strategies Affect Reading Comprehension among Intermediate Learners of Mandarin Chinese

Hongyi Jia
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd
Part of the Other Languages, Societies, and Cultures Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/2820

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Identifying How Successful Vocabulary-Learning Strategies Affect Reading Comprehension among Intermediate Learners of Mandarin Chinese

Hongyi Jia

A thesis submitted to the faculty of Brigham Young University in partial fulfillment of the requirements for the degree of Master of Arts

Matthew B. Christensen, Chair
Dana Scott Bourgerie
J. Paul Warnick

Center for Language Studies
Brigham Young University
August 2011

Copyright © 2011 Hongyi Jia
All Rights Reserved
ABSTRACT

Identifying How Successful Vocabulary-Learning Strategies Affect Reading Comprehension among Intermediate Learners of Mandarin Chinese

Hongyi Jia
Center for Language Studies, BYU
Master of Arts

The Chinese language has become an increasingly important Asian language for American students and more and more of them have a desire to learn it. The problems and the difficulties that Chinese foreign language (CFL) learners have when they try to use their knowledge of new vocabulary in reading comprehension are apparent. Previous studies have described some aspects that may influence reading comprehension and have indicated that Chinese is a relatively difficult language for native speakers of American English, but there are few studies that have focused on improving CFL learners’ Chinese reading comprehension. This study investigates the effect of pre-learning vocabulary and contextualized word learning strategies on Chinese reading comprehension. The purpose of this study is to examine whether pre-learning vocabulary and contextualized word learning positively affect CFL learners’ reading comprehension. The results show that pre-learning vocabulary does not have a positive effect on Chinese reading comprehension and that although contextualized word learning positively affects CFL learners’ reading comprehension the effect is not significant. This study also gives relative analysis.

Keywords: reading comprehension, Chinese foreign language (CFL) learners, word learning strategy, contextualized word learning, pre-learning, and word list learning.
ACKNOWLEDGMENTS

I am deeply grateful for those professors who have taught and helped me. First and foremost, I would like to thank Dr. Matthew B. Christensen, my chair, for patiently helping me develop my research questions and experiment design and for reading and editing multiple drafts of my thesis. I would like to thank Dr. Dennis L. Eggett for helping me with statistics. I also appreciate that Dr. Dana Scott Bourgerie and Dr. J. Paul Warnick have kindly served as my thesis committee members and gave me many useful suggestions.

I am indebted to Andrew and Rachel Grover for their hours helping me correct the English mistakes in my drafts and for their helpful suggestions.

I would like to thank Benjamin Wilcox for helping me write the vocabulary list, the examples sentences, the English questions, and for designing the reading comprehension test. I am thankful for the raters and students who participated in this study. I am also thankful for my roommate and friends who helped me through my study and my thesis.

Finally, I really want to say how deeply grateful I am for my parents who provided me with good opportunities to learn in the world and gave me much more than I could hope to return.
# Table of Contents

List of Tables ................................................................................................................................. vii

Chapter One: Introduction ............................................................................................................. 1

  Statement of the Problem ........................................................................................................... 1
  Statement of Purpose .................................................................................................................. 2
  Overview of the Study ............................................................................................................... 2
  Significance of the Study ......................................................................................................... 3

Chapter Two: Review of Literature ............................................................................................... 5

  Language Category Research .................................................................................................... 5
  Reading Strategies Research ...................................................................................................... 9
    Factors That May Influence Reading Comprehension ........................................................... 9
    Reading Comprehension and Vocabulary Acquisition. ......................................................... 12
    Chinese Reading Strategies. .................................................................................................. 14
  Reading Comprehension Test and Experiment Research ...................................................... 16
  Rationale for this Study ........................................................................................................... 20
  Research Questions ................................................................................................................. 21

Chapter Three: Research Design ................................................................................................ 23

  Design Overview ..................................................................................................................... 23
  Participants ............................................................................................................................... 25
  Materials .................................................................................................................................. 34
  Instruments ............................................................................................................................... 36
  Procedures ............................................................................................................................... 38
  Scoring ..................................................................................................................................... 41
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Analysis</td>
<td>41</td>
</tr>
<tr>
<td>Research Question Number One – Pre-learning Discussion.</td>
<td>41</td>
</tr>
<tr>
<td>Research Question Number Two – Contextualized Word Learning Discussion.</td>
<td>42</td>
</tr>
<tr>
<td>Other Related Points of Discussion</td>
<td>43</td>
</tr>
<tr>
<td>Chapter Four: Results</td>
<td>45</td>
</tr>
<tr>
<td>Basic Data Analysis</td>
<td>45</td>
</tr>
<tr>
<td>Research Question One Analysis</td>
<td>56</td>
</tr>
<tr>
<td>Research Question Two Analysis</td>
<td>65</td>
</tr>
<tr>
<td>Other Factors Analysis</td>
<td>71</td>
</tr>
<tr>
<td>Summary</td>
<td>74</td>
</tr>
<tr>
<td>Chapter Five: Discussion and Conclusions</td>
<td>75</td>
</tr>
<tr>
<td>Overview of the Study</td>
<td>75</td>
</tr>
<tr>
<td>Discussion of the Results</td>
<td>76</td>
</tr>
<tr>
<td>Discussion of Basic Data Analysis</td>
<td>76</td>
</tr>
<tr>
<td>Discussion of Research Question One</td>
<td>78</td>
</tr>
<tr>
<td>Discussion of Research Question Two</td>
<td>80</td>
</tr>
<tr>
<td>Discussion of Other Factors.</td>
<td>84</td>
</tr>
<tr>
<td>Implications of the Results</td>
<td>86</td>
</tr>
<tr>
<td>Limitations of this Study</td>
<td>87</td>
</tr>
<tr>
<td>Recommendations for Future Research</td>
<td>87</td>
</tr>
<tr>
<td>References</td>
<td>91</td>
</tr>
<tr>
<td>Appendix A: Consent Form I</td>
<td>99</td>
</tr>
<tr>
<td>Appendix B: Consent Form II</td>
<td>101</td>
</tr>
<tr>
<td>Appendix C m̀dàjiè kāīchē 马大姐开车</td>
<td>103</td>
</tr>
</tbody>
</table>
Appendix D 马大姐开车 Questions

Appendix E 马大姐开车 Vocabulary List

Appendix F 马大姐开车 Vocabulary List

Appendix G 中国科举制度 Questions

Appendix H 中国科举制度 Vocabulary List

Appendix I 中国科举制度 Vocabulary List

Appendix J 看不见的爱 Questions

Appendix L 看不见的爱 Vocabulary List

Appendix M 看不见的爱 Vocabulary List

Appendix O Directions for the Test
List of Tables

Table 1 Groups’ Information for Different Semesters .................................................. 28
Table 2 Previous Experience of Students ........................................................................ 29
Table 3 Raters’ Chinese Background Information .......................................................... 33
Table 4 Vocabulary Learning Strategies for Each Group .................................................. 37
Table 5 Groups for Research Question One – Pre-Learning .......................................... 42
Table 6 Groups for Research Question Two – Contextualized Word Learning ............... 43
Table 7 Means of Question 1 .......................................................................................... 46
Table 8 P value of t-test for Question 1 ........................................................................... 47
Table 9 Means of Question 2 .......................................................................................... 47
Table 10 P value of t-test for Question 2 ....................................................................... 48
Table 11 Means of Question 3 ........................................................................................ 48
Table 12 P value of t-test for Question 3 ....................................................................... 49
Table 13 Means of Question 4 ........................................................................................ 50
Table 14 P value of t-test for Question 4 ....................................................................... 50
Table 15 Means of Question 5 ........................................................................................ 51
Table 16 P value of t-test for Question 5 ....................................................................... 52
Table 17 Means of Total Score ...................................................................................... 53
Table 18 P value of t-test for Total Score ...................................................................... 53
Table 19 Groups of Highest and Lowest Means for Each Question and Total Score ....... 54
Table 20 T-test of Question 1 for Analyzing Pre-Learning .............................................. 56
Table 21 T-test of Question 2 for Analyzing Pre-Learning .............................................. 58
Table 22 $T$-test of Question 3 for Analyzing Pre-Learning .............................................................. 59
Table 23 $T$-test of Question 4 for Analyzing Pre-Learning .............................................................. 59
Table 24 $T$-test of Question 5 for Analyzing Pre-Learning .............................................................. 60
Table 25 $T$-test of Total Score for Analyzing Pre-Learning ............................................................. 61
Table 26 Groups of Highest and Lowest Means for Each Question and Total Score for Pre-Learning .................................................................................................................................. 62
Table 27 Significantly Different Groups for Each Question and Total Score for Pre-Learning .. 63
Table 28 $T$-test of Question 1 for Analyzing the Contextualized Word Learning.......................... 65
Table 29 $T$-test of Question 2 for Analyzing the Contextualized Word Learning.......................... 66
Table 30 $T$-test of Question 3 for Analyzing the Contextualized Word Learning.......................... 67
Table 31 $T$-test of Question 4 for Analyzing the Contextualized Word Learning.......................... 67
Table 32 $T$-test of Question 5 for Analyzing the Contextualized Word Learning.......................... 68
Table 33 $T$-test of Total Score for Analyzing the Contextualized Word Learning ...................... 68
Table 34 Groups of Highest and Lowest Means for Each Question and Total Score for Contextualized Word Learning ........................................................................................................ 69
Table 35 $P$ Values of the Model ........................................................................................................ 71
Table 36 Means of Different Raters.................................................................................................... 72
Table 37 Pearson Correlation Coefficients of Raters for Question 1 .............................................. 72
Table 38 Pearson Correlation Coefficients of Raters for Question 2 .............................................. 72
Table 39 Pearson Correlation Coefficients of Raters for Question 3 .............................................. 73
Table 40 Pearson Correlation Coefficients of Raters for Question 4 .............................................. 73
Table 41 Pearson Correlation Coefficients of Raters for Question 5 .............................................. 73
Chapter One: Introduction

Statement of the Problem

After teaching Chinese at Brigham Young University as a teaching assistant, I found that one difficulty American adults have when learning Chinese is using their knowledge of new vocabulary for reading comprehension. They may know the translations of new words that they are learning, but they still have difficulty understanding a reading passage that includes the new words. Therefore, we think finding good vocabulary-learning strategies that affect reading comprehension for Chinese Foreign Language (CFL) is a meaningful and interesting study.

Previous research in the field of teaching Chinese indicated that Chinese is a relatively difficult language for native speakers of American English to learn for three reasons: 1) the language family; 2) the writing system; 3) the unique characteristics of Chinese that are significantly different from English. Previous studies on improving reading comprehension described the factors that may influence reading comprehension as vocabulary knowledge (Stratton, & Nacke, 1974) and context (Mezynski, 1983; Stahl, 1990), which implies that learning vocabulary with contextual information may have a positive effect on reading comprehension. Previous research also discussed if pre-learning target words, which is to study some target words before reading, improves reading comprehension. Previous studies on Chinese reading comprehension concluded that one problem of a poor Chinese reader was that they could not contextually understand the texts without prior knowledge of vocabulary (Lau, 2006).
Statement of Purpose

The primary purpose of this thesis is to determine whether the two vocabulary learning strategies of pre-learning vocabulary and contextualized word learning, which is to study new words with example sentences, have positive effects on CFL learners’ reading comprehension. This study attempts to illuminate the possible reasons behind the effects of the two strategies on the participants’ reading comprehension in this experiment. This study also analyzes other factors that may influence the effects the two vocabulary learning strategies have on reading comprehension in this experiment and the limitation of this experiment. Specifically, this study addresses the following research questions:

1) Does pre-learning vocabulary instruction affect CFL (Chinese Foreign Language) learners’ reading comprehension?
2) Does teaching contextualized word learning strategy affect CFL learners’ reading comprehension?

Overview of the Study

This study designed six treatment groups using six treatments that are combinations of two or three of the four different vocabulary learning strategies respectively: 1) contextualized word learning, which is reading a vocabulary list with example sentences; 2) vocabulary list learning, which is reading a vocabulary list without example sentences; 3) pre-learning, which is reading a vocabulary list before reading a passage; and 4) normal-learning, which is reading a vocabulary list at the same time as reading a passage. The results groups are:

Group 1: Control
Group 2: Contextualized word learning and pre-learning
Group 3: Vocabulary list learning and pre-learning
Group 4: Contextualized word learning, pre-learning, and normal learning

Group 5: Vocabulary list learning, pre-learning, and normal learning

Group 6: Contextualized word learning and normal learning

Group 7: Vocabulary list learning and normal learning

In addition to these groups we also have combined groups which are collections of the last six groups with an overlapping learning technique used. Each of these combined groups is named after the groups that it is composed of. The hybrid groups are:

Group 2-4-6: Contextualized word learning

Group 3-5-7: Vocabulary list learning

Group 2-3: Pre-learning

Group 4-5: Pre-learning and normal learning

Group 6-7: Normal learning

Three raters evaluated each answer sheet. The study used SAS statistics software to calculate mean and standard error of each question and the total scores for the seven groups and to analyze variance and comparison of different groups according to the two research questions by using the average of the three raters’ evaluations.

**Significance of the Study**

This study attempts to suggest a good way to apply appropriate reading strategies to reading Chinese, which most previous studies have not focused on.

The results of this study show significant gains made in pre-learning with normal learning, while pre-learning without normal-learning is actually worse than normal learning. The results of this study also show that the contextualized word learning strategy has a positive effect on reading comprehension; however, it is not always statistically significant. These results
suggest that long-term training for learning contextualized word learning strategy and pre-learning may be good for the learners, especially for intermediate students.
Chapter Two: Review of Literature

The purpose of this study is to determine whether pre-learning vocabulary instruction and contextualized word learning strategy affect Chinese reading comprehension for intermediate Chinese foreign language learners. The overall objective of this chapter is to provide a review of previous studies that focus on improving reading comprehension and the effects of vocabulary learning on reading comprehension. This chapter starts with reviewing language category research, which explains that not only is Chinese a relatively difficult language for native speakers of American English, but also that context is very important for Chinese reading comprehension. Next, it discusses reading strategy research, which explains which factors may influence reading comprehension, the relation between reading comprehension and vocabulary acquisition, and Chinese reading strategies. In addition, this chapter discusses reading comprehension experimental research. After reviewing these studies, implications for this study describe why it is worthwhile to study and test methods of learning Chinese vocabulary in order to improve English native learners’ Chinese reading comprehension. Finally, it proposes the research questions for the current study.

Language Category Research

Chinese is a relatively difficult language for native speakers of American English to learn. Regarding the length of time that learners need to spend in order to attain a certain high-proficiency level, according to the categorization used by the Defense Language Institute Foreign Language Center, Chinese falls into the most difficult of four possible categories:

“Category I: English, French, Italian, Portuguese (Brazilian), Portuguese (European), and Spanish; Category II: German, Romanian (DLPT III); Category III: Czech, Greek, Hebrew,
Persian-Farsi, Polish, Russian, Serbian/Croatian, Tagalog, Thai, Turkish, Ukrainian, and Vietnamese; Category IV: Arabic, Chinese Mandarin, Japanese, and Korean” (DLIFLC document; Everson, 1994; McEwen, 2006).

For languages in Category IV, such as Chinese, 1,320 contact hours are required in order to attain the second proficiency level, while only 480 contact hours are required for learning Category I languages to be the same proficiency level (Christensen & Warnick, 2006; Everson, 1994; McEwen, 2006). From the Defense Language Institute Foreign Language Center Catalog (2009-2010) (2009), we can see that the course for basic Chinese-Mandarin needs sixty-four training weeks, while the course of basic Spanish only needs twenty-six training weeks.

There are three aspects of the Chinese language that cause difficulties for Native American English speakers: the language family, the writing system, and the unique characteristics of the Chinese language that are significantly different from English.

Considering evolvement, history, and comparisons of languages, according to the genetic classification of languages, English belongs to the Germanic language family (Fromkin, Rodman, & Hyams, 2007), and is close to French, Italian, Portuguese, and Spanish, which belong to the Romance (Latin) language family (Fromkin et al., 2007). As the web of The Sino-Tibetan Family of Languages (2000) said, both Germanic languages and Romance (Latin) languages are in the Indo-European language family whose verbs and nouns have different endings because of different situations. The relation between English and Romance (Latin) languages, such as French and Spanish, is like that of cousins sharing family bloodlines (Fromkin et al.).

However, English is far from Chinese, which belongs to the Sino-Tibetan language family that contains tonal languages (DeFrancis, 1984; Sino-Tibetan language Wiki, 2010).

“There are two kinds of phonemes in Chinese: segmental phonemes, which may be thought of as
sequential sounds, and suprasegmental phones or tones, which in a sense are added to the syllable as a whole” (Defrancis, p. 42). “Suprasegmental phonemes or tones, which give Chinese speech its distinctive musical or singsong quality, must be distinguished from intonation” (Defrancis, p. 46). “… the Chinese syllable is distinctive in that, in most cases, it constitutes a morpheme, the smallest unit of meaning” (Defrancis, p. 46). These characteristics are distinct from those of English. Thus, English speakers learning Chinese are confronted with a language that does not have any connection with their own language family.

Considering the history of the writing system, English uses alphabetic writing that relates to sound, and is a Romanized language, which is derived from the Greek system (Fromkin et al.). Chinese is logographic or word writing (Fromkin et al.), which is derived from pictograms and ideograms: “logograms, the symbols of a word-writing system, are ideograms that represent in addition to the concept, the word or morpheme in the language for that concept” (Fromkin et al. p. 509). This word writing system that does not have much connection with the sound system can be awkward for English and other Indo-European languages in which most words have more than one morpheme, which relates to how to write them (Fromkin et al.). This is another reason that American learners have difficulties learning Chinese.

Analyzing the characteristics of Chinese may provide a detailed explanation for why Chinese is a relatively difficult language for native speakers of American English to learn. There are three main characteristics of Chinese that are significantly different from English. First, each Chinese character represents an individual syllable or morpheme and has its own meaning, which may change depending on different contexts. “Because most syllables have meaning they are often considered to be words” (DeFrancis, p. 46). Longer Chinese words can be made by combining two or more characters. However, in English most words are compounded by more
than one morpheme, and the meaning changes according to changes in morphemes (Fromkin et al.).

Second, Chinese characters do not provide many clues that indicate pronunciation (Everson, 1994). “In Chinese one syllable may correspond to several different morphemes” (Everson, p. 543). On the other hand in alphabetic writing, such as English, the writing system has a close relation to the sound system. “The Chinese characters are not based on any inventory of symbols comparable to the alphabetic scripts used to represent English, Russian, Arabic, and other languages” (DeFrancis, p. 91). “Chinese chose a semantic basis rather than a phonetic one for their system of classification” (DeFrancis, p. 93).

Although Chinese has purely semantic characters and radicals plus phonetic characters (DeFrancis, 1984), and it is possible that Chinese characters can be guessed by phonetic elements, but it is nearly impossible for unskilled readers. “Ability to do so will of course vary widely depending on the endowments and training of the reader. … Guessing the pronunciation and meaning of characters when reading has been honed to a fine skill among users of character-based scripts” (DeFrancis, p. 130). Thus, skilled readers who have knowledge of understanding through the context can learn or get to know some new words in the course of reading.

Third, Chinese does not have inflection that can change the meaning of a noun or express different statuses of a verb in different situations as in English. However, English has pervasive inflectional morphemes, which label tense, voice, person, number, and so on (Fromkin et al.). Two characteristics of Chinese nouns are that: “they are mass nouns comparable to the English mass noun ‘rice.’ …they do not undergo change to distinguish singular and plural” (DeFrancis, p.47). Also, “Chinese is characterized by having aspect rather than tense. This technical term refers to the way a speaker looks at an event or state” (DeFrancis, p. 48). “In contrast to English
adjectives, … they incorporate the idea which we represent by words like ‘am’, ‘is’, ‘are’,…” (DeFrancis, p. 48). We need the context to know if there is more than one person or if something happened. Therefore, we can see that context is very important for us to know a word’s exact meaning. “In contrast to English, which is a sentence-oriented language, what might be considered omissions or deletions are much more common in Chinese, which is context-oriented” (DeFrancis, p. 50).

Through analyzing the characteristics of the Chinese language, we can know not only that Chinese is a relatively difficult language for native speakers of American English, but also that understanding a Chinese character or word depends more on context, and understanding through context is very important for learning Chinese characters and improving Chinese reading comprehension.

**Reading Strategies Research**

In order to improve reading comprehension, we first need to know what factors can affect reading comprehension and why some learners are poor readers.

**Factors That May Influence Reading Comprehension.** Stuart Webb (2009) investigated the effects of pre-learning vocabulary on reading comprehension and writing. After reviewing the previous studies, Webb (2009) summarized previous studies that vocabulary knowledge (Stratton, & Nacke, 1974), prior background knowledge about the given topic (Stahl, Hare, Sinatra, & Gregory, 1991; Stahl, & Jacobson, 1986; Stahl, Jacobson, Davis, & Davis, 1989), context (Mezynski, 1983; Stahl, 1990), and the density of unknown words (Mezynski, 1983; Pany & Jenkins, 1978) influence reading comprehension. When readers have more vocabulary knowledge, more prior background knowledge about the given topic, more abilities
of understanding through the context, and fewer unknown words, they may have more understanding of texts.

Webb (2009) also summarized previous studies that vocabulary factors—such as length (Baker, 1989), conceptual complexity, polysemy, and homonymy (Stahl, 1990)—can influence reading comprehension. “longer words are more likely to disrupt understanding of a passage than shorter words” (Webb, p. 446). “One might similarly assume that words at a high level of conceptual difficulty would more strongly affect comprehension, because they would be more likely to carry a higher percentage of the content of the passage than less conceptually complex words (Stahl, p. 12)” (as cited in Webb). Although we are not sure if polysemy or homonymy create genuine comprehension difficulties, “context could (and certainly does) disambiguate polysemous words (Stahl, p. 12)” (as cited in Webb). Especially, Webb concluded Stahl’s study that some scholars think all words are polysemous (Stahl, 1990).

From these studies we can see that the main factors that affect reading comprehension are vocabulary knowledge and the ability to understand through context, as all other elements relate to these two factors. We can also analyze the characteristics of poor and good readers and find what can help learners improve their reading comprehension.

Lau (2006) compared the difference between good and poor readers, and she reviewed previous studies and summarized the following ideas:

Poor readers mainly have problems in using the following strategies: (1) they do not know how to construct the main ideas and macrostructure of the texts (Brown & Day, 1983; Day & Zajakoski, 1991; Williams, 1986); (2) they are not familiar with the text structure and do not make use of the text structure to organize the main ideas (McGee, 1982; Meyer, Brandt & Bluth, 1980); (3) they have little prior knowledge and do not
know how to activate their knowledge to facilitate text comprehension (Afflerbach, 1990; Stahl, Jacobson, Davis & Davis, 1989); (4) they have difficulties in drawing inferences to achieve in-depth understanding of the texts (Long, Seely, Oppy & Golding, 1996; Singer, 1994); (5) they lack metacognitive ability, and are not aware of the problems that emerge during reading (Duffy et al., 1987; Garner, 1987) and do not know how to monitor their reading process (Baker & Brown, 1984; Garner, 1987). (p. 384)

Stahle (1990) also summarized the differences between good and poor readers in four aspects: prior word knowledge, topic knowledge, verbal ability, and working memory. When readers know more words, they can learn more words, because in context they can get richer contextual information about words that they do not know (Stahle, 1990).

In short, poor readers cannot contextually understand the texts without prior vocabulary knowledge. Thus, we see that contextually understanding vocabulary has a significant impact on reading comprehension.

The key reason that readers do not have the ability to contextually understand the texts is that unsuccessful readers read texts in a word-by-word way. Singhal (1998) reviewed Block’s study and mentioned that unsuccessful readers do not direct attention toward the text (Block, 1986). Bernhardt’s study (as cited in Everson, 1994) suggested the use of text-based and extra-text-based components. Text-based components include word recognition, and extra-text-based components include intra-textual perception (as cited in Everson, 1994). Nassaji (2002) also wrote, “skilled L2 readers, who possess the necessary syntactic and semantic knowledge, may still lack adequate top-down strategies to sample the text rapidly, or may simply fail to use their conceptual knowledge when reading for comprehension” (p. 458).
We can see that having the ability of word recognition and intra-textual perception are two important factors for reading comprehension. Considering the characteristics of Chinese mentioned above, having the ability of word recognition is the foundation for having the ability of intra-textual perception in Chinese reading. However, even if readers have the ability of word recognition, which means only knowing the English translation, it does not mean they have the ability of intra-textual perception. The ability of word recognition means L2 readers know the translation of the word, but it does not mean they have the contextual knowledge of the word. The contextual knowledge of a word includes using the word in an appropriate way and understanding the different meanings of a Chinese word in different situations.

In order to improve Chinese learners’ reading comprehension, we need to help them find a good way to learn Chinese words and to use contextual knowledge of the target Chinese words to read the text contextually. We think that learning target words in meaningful sentences is a good way to understand the contextual knowledge of a word. Having the knowledge of words can help readers to read the Chinese text contextually. The current study tries to show this. Below we review what previous studies have accomplished in relation to vocabulary acquisition and reading comprehension, which is important to this study.

**Reading Comprehension and Vocabulary Acquisition.** Susanne Rott (2004) studied the relationship between text comprehension and word retention through testing the effects of different tasks, including “a cued-output task (CO),” “a self-selected-out-put task (SO),” and “a reading-only control task (RO).” Rott reviewed Parry and Kintsch’s studies that showed how using text-based and learner-based information L2 readers could understand the context without the knowledge of every single word in the passage (Parry, 1993, 1997; & Kintsch, 1998). Rott also stated her own opinion in support of previous studies, which state that vocabulary
acquisition has a positive effect on reading comprehension, after he reviewed the studies of Pak; Holly & King; Jacobs, Dufon, & Hong; Davis; Rott, Williams, & Cameron (Pak, 1986; Holly & King, 1971; Jacobs, Dufon, & Hong, 1994; Davis, 1989; Rott, Williams, & Cameron, 2002).

Rott organized an experiment in which consisted of forty-seven English native volunteers who were German learners. There were three reading conditions as treatments: a cued-output task (CO), a self-selected-out-put task (SO), and a reading-only control task (RO). The experiment was to test a vocabulary pre-test, a vocabulary post-test, and comprehension. During the fifth, sixth, and seventh weeks the reading treatments were given to the participants. Through analyzing the testing results, Rott concluded that word acquisition had a strong relation to text comprehension. Rott also said, “It appears that when L2 learners received word-focused activities in addition to a text, learning took place independently from passage comprehension, a finding also reported in Rott et al. (2002)” (Rott, p. 193).

Webb (2009) concluded and cited the studies of Beck, Perfetti, & McKeown and McKeown, Beck, Omanson, & Perfetti that contextual and definitional learning can improve reading comprehension (Beck, Perfetti, & McKeown, 1982; McKeown, Beck, Omanson, & Perfetti, 1983). Also, he concluded Mezynski’s study that “using a variety of tasks to develop a deeper knowledge of target words” (Webb, pp. 443) can help to improve reading comprehension (Mezynski, 1983). However, only studying target words and their definition could not significantly improve reading comprehension, which is what Webb concluded Stahl & Fairbanks’ study (Stahl, & Fairbanks, 1986).

These previous studies support the idea that studying the effect of pre-learning vocabulary and contextualized word learning is a worthy endeavor in order to improve reading
comprehension. The aims of the current study is to improve Chinese reading comprehension, so it is necessary to review previous articles that focus on Chinese reading problems and strategies.

**Chinese Reading Strategies.** Lau (2006) concluded that poor readers of Chinese have the problems mentioned above—poor readers cannot contextually understand the texts without prior knowledge. However, there is difference between native Chinese readers and L2 Chinese readers whose native language is English. Hayes’ study (as cited in Ke, 1996) found that native Chinese readers use a phonetically oriented strategy while nonnative readers are uncertain whether to use a phonological or graphic strategy, and “the nonnatives made more graphic errors at the sentence level than did the natives” (Ke, 1996, p. 341). Hayes’ study (as cited in Everson, 1994) found that American readers tend to use pictorial clues to recognize Chinese characters that lead them to make errors and take a long time to read. Bernhardt’s model (as cited in Everson, 1994) predicted that the inability to recognize Chinese characters quickly and accurately leads readers to have reading comprehension problems.

These studies show that native English speaking Americans who are learning Chinese have difficulties in recognizing Chinese characters in the reading process. This is because nonnative readers do not know the characteristics of Chinese well and stereotype Chinese, for example reading a Chinese character only using a radical. Sometimes they try to guess Chinese vocabulary with their limited knowledge, which leads them to make mistakes (DeFrancis, 1984). As they lack the ability to recognize words, it is difficult for them to have any intra-textual perception, a topic mentioned above. In order to improve nonnative Chinese reader’s reading comprehension, we need to help them grasp strategies of recognizing the meaning of Chinese characters and understanding Chinese sentences so that they can comprehend a Chinese text contextually.
Wen (1998) gave some reading strategies through analyzing the characteristics of Chinese characters and Chinese language.

In order to recognize the meaning of Chinese characters, readers (learners) need to know the information about the structures and compounds of the characters. Chinese characters can be deconstructed by analyzing their components and radicals. Every radical has a cue, although some of them are not obvious (Wen, 1998).

In order to understand a Chinese text, there are three strategies to reading Chinese sentences: shortening the sentence, looking for the structural words, and looking for contextual clues. These strategies are helpful because compared with English the Chinese language has the following characteristics:

“English is a syntactically based language, and the subject-predicate is its basic structure. In comparison, Chinese language is topic-prominent. … The structure of Chinese sentence is flexible” (Wen, p. 7).

“Tsao (1979) states that Chinese language is discourse-oriented and the topic in a discourse plays an important role. … Therefore, when reading Chinese, it is important to read the whole paragraph rather than a sentence or a clause” (Wen, p. 7).

Wen also said, “They [some readers] tend to focus on vocabularies and surface structures and fail to grasp the main ideas of the macro-structure” (Wen, p. 12).

Native speakers of American English who are learning Chinese have difficulty contextually understanding a Chinese text because they have difficulty recognizing the meanings of Chinese characters, which includes the translation of the word and the contextual knowledge of the word. After analyzing Chinese characters, we can see that in order to improve Chinese reading comprehension it is more important than in other languages not to read Chinese texts in a
word-for-word way. In order to have the ability to look for contextual clues in a Chinese text, we first need to solve the problem of recognizing the meanings of Chinese characters, which includes knowing the translation of the word and having the contextual knowledge of the word.

In short, some native speakers of American English who are learning Chinese do not have enough cultural knowledge of how Chinese characters are formed to learn, remember, and recognize the meaning of Chinese characters. Using meaningful sentences to help these learners have vocabulary knowledge before they read a text may be a good way to help them contextually understand Chinese text and improve their reading comprehension ability, because reading new words in a sentence can help readers understand these new words within a target situation. This is especially effective if the meaning of the new words in the sentence is the same as the meaning that is used in the text. A sentence that uses the new words can explain the meaning of the new words through example.

**Reading Comprehension Test and Experiment Research**

How to measure the improvement of reading comprehension with using meaningful sentences is another point that this study should consider. Below we review some studies of testing reading comprehension.

In general, there are three purposes to test reading comprehension: first is to sort students by their scores, second is to diagnose learners’ reading problems, and third is to evaluate if it is effective to apply an experimental treatment (Aweiss, 1993; Cross, & Paris, 1987).

The current study focuses on the last two purposes, as I attempt to show that pre-learning vocabulary and contextualized word learning are good ways to improve Chinese learners’ reading comprehension. This is specifically responding to American English natives’ reading
problem: having difficulties in recognizing Chinese characters in the reading process, and having difficulty in using vocabulary knowledge to contextually understand a Chinese text.

As mentioned below, from the previous studies Aweiss (1993) summarized four methods to test reading comprehension and discussed the advantages and disadvantages of these methods.

The first method to test reading comprehension is a multiple-choice test. The disadvantages are that the correct answers can be obtained by guessing (Bernhardt, 1991; Henning, 1987; Pyrczak, 1975; Klein-Braley, 1985) and the results of the tests may not truly reflect reading comprehension (Klein-Braley, 1983, 1985, & 1991).

The second method is a short answer test. The advantage to this method is that it allows learners to express themselves freely. The disadvantage is that learners (readers) need to write; therefore, learners’ writing ability can impact the testing results (Weir, 1990).

The third method is a cloze test. The advantage is that it is better than a multiple-choice test, as it can reflect more of the text-based reading comprehension process (Swaffer, Arens, & Byrnes, 1991). The disadvantage is that the testing results can also be influenced by learners’ other language abilities.

The fourth method is the recall protocol test. The advantage to this test is that it is the most straightforward test and provides the most valid information of any other reading comprehension test (Johnson, 1983). The disadvantage is that learners (readers) can write without an objective weighting analysis (Maria, 1990; Page, 1990; Swaffer, et al., 1991). Learners (readers) can write what they remembered without understanding.

Chang (2006) who studied immediate recall task and second language reading comprehension, and Pulido (2007), who studied the relationship between text comprehension and second language vocabulary acquisition mentioned the effect of memory on the recall task.
The recall task can measure not only reading comprehension but also memory. “The findings that the immediate recall task measuring both memory and understanding generates an incomplete picture of readers’ comprehension have implications for the assessment of L2 reading comprehension” (Chang, 2006, p. 537).

Since the current study focuses on measuring the effect of contextually understanding target words on reading comprehension of a given passage, we think a short answer test is appropriate. The writing ability of learners (readers) may impact the testing results; however, we can decrease the possibility of the influence of their writing ability through designing questions in English and asking them to answer in English. Also, English questions do not support any clue that can help learners (readers) answer the questions without real understanding. The current study tries to decrease the influence of other factors, such as readers’ memory, readers’ guessing ability, and so on.

There are not many studies investigating Chinese reading strategies and Chinese reading comprehension, and since the characteristics of Chinese language are totally different from Romanized languages, such as English. There are only a few studies that can be referenced to design the experiment that tests the effect of pre-learning vocabulary and contextualized word learning on Chinese reading comprehension.

Paul (2005) tried to identify and evaluate what reading strategies are effective for intermediate and advanced level non-native readers of Chinese when they read a short story. Although Paul used think-aloud protocols and a verbal recall protocol to measure the effect of the reading strategies, which we did not choose, his conclusion based on his experiment enlightened me. “This suggests that how a subject uses reading strategies may be more important than which ones she uses” (Paul, p. 115). “Instruction given to advanced level readers should
focus on native-like top-down processing strategies” (Paul, p. 118). Reading strategies at a high level should focus on how to understand the meaning of the whole text, which requires the contextual understanding of important words.

Webb’s (2009) study designed an experiment to test the effects of pre-learning L2 vocabulary on reading comprehension and writing. The participants were 71 Japanese native speakers who were studying English in a first-year class at a university in Japan. They were divided into two groups. One group used a receptive learning method, and the other one used a productive learning method. First, both groups studied fifteen English words with translations in a receptive method. However, the receptive treatment group was instructed to cover the translation and try to recall each word’s translation when they learned the vocabulary while the productive treatment group did not do this. After studying the vocabulary, both groups had a picture-description test, a reading-comprehension test, and a translation test.

The results showed that pre-learning vocabulary improved reading comprehension and that vocabulary knowledge may help comprehension. It also showed that receptive learning, “such as looking up words in a dictionary, matching words with their meanings or definitions, guessing from context, and learning from word pairs” (Webb, 2005, p. 34), is better than productive learning, “such as cloze exercises or writing tasks” (Webb, p. 34), which implies that contextualized word learning may be effective.

Although Webb’s (2009) study tested English reading comprehension, and it measured sentence comprehension, his experiment is still a good reference for designing an experiment to test the effects of pre-learning vocabulary on Chinese reading comprehension.

We also read articles that were written in Chinese and stated contextual understanding is important to learning Chinese and reading comprehension. Tan’s (1997) study supported that
background information is one important factor that influence reading comprehension. Yuan’s (1998) study expressed that learning new vocabulary through reading comprehension exercises and enlarging background knowledge is good for improving Chinese foreign language learners’ reading comprehension. Zhu (2004) studied Chinese language teaching in America. He found that in America teaching methods focused on applying the information in practice and teaching new vocabulary in context.

Rationale for this Study

As we reviewed the studies above, there are some studies that focused on improving reading comprehension and the effects of vocabulary learning on reading comprehension, but there are few studies that focused on improving Chinese foreign language learners’ Chinese reading comprehension. The articles that studied Chinese reading comprehension analyzed the reasons that contextualized word learning is good for Chinese reading comprehension, the way that the meaning of Chinese words depends more on the context than other languages’ words, the fact that “Chinese language is topic-prominent” (Wen, p. 7), and tried to find good reading strategies for Chinese reading comprehension. However, these previous studies did not suggest a good way to apply appropriate reading strategies to reading Chinese.

This study investigates the effect of pre-learning Chinese vocabulary and using example sentences to learn Chinese vocabulary on Chinese reading comprehension. Pre-learning Chinese vocabulary is a way to help Chinese foreign language learners solve one of their most serious reading problems—the difficulty they have recognizing Chinese characters when they are reading Chinese passages. Therefore pre-learning Chinese vocabulary can help them apply top-down processing strategies. Using meaningful sentences to learn Chinese vocabulary represents
contextualized word learning that corresponds to the characteristics of Chinese language, which can help foreign language learners improve their Chinese reading comprehension.

**Research Questions**

The purpose of this study is to test whether pre-learning vocabulary and contextualized word learning have a positive effect on improving Chinese reading comprehension. This study will examine the following research questions:

1) Does pre-learning vocabulary instruction affect CFL (Chinese foreign language) learners’ reading comprehension?

2) Does teaching a contextualized word learning strategy affect CFL learners’ reading comprehension?
Chapter Three: Research Design

The aim of this study is to analyze the effects of two vocabulary learning strategies on CFL learners’ reading comprehension. This study also attempts to find the possible reasons that explain the effect of these two vocabulary learning strategies. This chapter explains the design of the experiment based on the aim, which has seven subsections: Design Overview, Subjects, Materials, Instruments, Procedures, Scoring, and Data Analysis. The Design Overview section describes the design of the experiment and states this study’s two hypotheses. The Subjects section introduces participants including the subjects and raters. The Materials section introduces what reading materials were used to test reading comprehension, what vocabulary words were chosen to use as the target words, what meaningful sentences were chosen in contextualized word learning treatment, and what questions were chosen in the tests. The Instruments section introduces the six treatments that were used. The Procedures section describes the process of the tests. The Scoring section states how the subjects’ answers were converted into scores. The Data Analysis section introduces the methods of analyzing the data of this experiment.

Design Overview

This experiment is designed to test the effect of different vocabulary learning strategies on text reading comprehension. There are seven groups. Group 1 is tested without any treatment, which is the control group. There are four different vocabulary learning strategies: contextualized word learning, vocabulary list learning, pre-learning, and normal-learning (at the same time as reading a passage). Thus, we designed six treatments groups that combined two or three of these four vocabulary learning strategies. Group 2 reads a vocabulary list with example sentences before participants get a reading passage, which represents combining contextualized
word learning and pre-learning. **Group 3** reads a vocabulary list before participants get a reading passage, which represents combining vocabulary list learning and pre-learning. **Group 4** reads a vocabulary list with example sentences before participants get a reading passage, and participants are allowed to refer to the list and sentences while reading the passage. This represents combining contextualized word learning, pre-learning, and normal learning. **Group 5** reads a vocabulary list before participants get a reading passage and they are allowed to refer to the list while reading the passage, which represents combining vocabulary list learning, pre-learning, and normal-learning. **Group 6** is given a vocabulary list with example sentences at the same time as a reading passage, which represents combining contextualized word learning and normal-learning. Finally, **Group 7** is given a vocabulary list at the same time as a reading passage, which represents vocabulary list learning and normal-learning.

We created **Group 2-4-6** by combining Group 2, Group 4, and Group 6 to represent contextualized word learning. I also created **Group 3-5-7** by combining Group 3, Group 5, and Group 7 to represent vocabulary list learning. Comparing **Group 2-4-6** and **Group 3-5-7**, we can see the effect of the contextualized word learning.

We created **Group 2-3** by combining Group 2 and Group 3, to represent pre-learning; **Group 4-5** by combining Group 4 and Group 5 to represent pre-learning and normal-learning; and **Group 6-7** by combining Group 6 and Group 7 to represent normal-learning. The differences between **Group 2-3**, **Group 4-5**, and **Group 6-7** can be used to analyze the effect of pre-learning.

We used Statistical Analysis Software (SAS) to calculate the means and analyze the variance for each group on each question and the total raw score. Therefore, we can see a comparison and analyze the differences between these groups.
There are two hypotheses:

1) There is a positive effect from pre-learning vocabulary instruction on CFL learners’ reading comprehension.

2) There is a positive effect from contextualized word learning strategies on CFL learners’ reading comprehension.

**Participants**

The subjects who participated in this study were Chinese language learners enrolled in the Chinese language 301 classes in the Fall Semester of 2009, the Winter Semester of 2010, the Fall Semester of 2010, and the Winter Semester 2011 at Brigham Young University, located in Provo, Utah.

Chinese 301 is designed to help students develop all four Chinese skills of listening, speaking, reading, and writing, and the course has emphasis on the spoken and written language at the intermediate level. This course aims to help students begin to read authentic Chinese texts. Students of Chinese 301 have five fifty-minute class periods each week. Generally, Chinese 301 courses are taught by one native Chinese instructor or one native Chinese instructor and one Chinese teaching assistant. Both the instructor and the teaching assistant teach grammar from the textbook, and they do not use ACT/FACT model. As explained by, the developers of this methodology, “The FACT classes include discussion about the language and the culture, … The ACT classes provide opportunities for learners to perform in the culture” (Christensen & Warnick, 2006, p. 59).

Normally students in the 301 class have completed four formal semesters of Chinese at the university, making the 301 class the fifth semester of their Chinese study. However, at BYU most of the students in the 301 class have had the previous experience of serving as full-time
Chinese-speaking missionaries for The Church of Jesus Christ of Latter-day Saints, and these students did not have all of the four formal semesters of Chinese study. Many of these students begin their formal Chinese language training with the 3rd or 4th semester course.

Students at this level were chosen because they can have certain abilities to read and do reading comprehension exercises, while in the first four semesters students have limited knowledge of vocabulary making it difficult for them to be involved in the type of reading comprehension that needs comprehensive abilities. Chinese 301 students are ideal for this study because they are being trained to read authentic Chinese texts, which require reading comprehension skills. Another reason for choosing 301 students is that, in general, the Chinese 301 course has two sections and around forty students every semester. So we can use only Chinese 301 students and do not need to combine different classes, which is good for us to make sure that participants are at a relatively consistent level.

During the Fall Semesters of 2009 and 2010 and the Winter Semester of 2011, I was the teaching assistant for Chinese 301 and was responsible for students’ reading comprehension. Thus, I had students do regular reading comprehension exercises in a normal class setting as the experiment was processing. After the students finished the reading exercises, the researcher gave them the consent form (see Appendix A) and explained to them the nature of the study. If they consented that the researcher could analyze their results of the reading comprehension exercises for the study, they signed the form and would get an extra ten points for the class. If they did not consent, they could do another culture activity to get the same extra ten points.

During the Winter Semester of 2010, the researcher contacted other teaching assistants who could contact students of Chinese 301. If the students wanted to help the researcher with the experiment, they could make an appointment and do the exercises after class. When they came,
the researcher gave them the consent form (see Appendix B). If they signed the form, they could get extra help for their Chinese class.

There were one hundred twenty-four subjects that did the reading comprehension exercises and signed the consent form. Twenty-one were female. One hundred twenty-one were Americans, two subjects were Korean, and one was Singaporean. The three who are not Americans were pursuing their bachelor’s degrees at BYU. All the students were divided into seven groups randomly by teaching assistants who did not know the students at all. **Group 1:** No treatment given. **Group 2:** Five minutes were given for students to read a vocabulary list with example sentences. When the time was up the vocabulary lists were collected and students were given a reading passage. **Group 3:** Five minutes were given for students to read a vocabulary list. When time was up the vocabulary lists were collected and students were given a reading passage. **Group 4:** Five minutes were given for students to read a vocabulary list with example sentences. Students were then given a reading passage and were allowed to refer to the vocabulary list with example sentences as they read. **Group 5:** Five minutes were given for students to read a vocabulary list. Students were then given a reading passage and were allowed to refer to the vocabulary list as they read. **Group 6:** Students were given a vocabulary list with example sentences and a reading passage at the same time. They were allowed to refer to the vocabulary list with example sentences as they read. **Group 7:** Students were given a vocabulary list and a reading passage at the same time. They were allowed to refer to the vocabulary list as they read.

As shown in Table 1, in the Fall Semester of 2009, there were thirty-two students who did the reading comprehension test, only Group 4 and Group 5 had four students, and the other groups had five students. There were six female students who were in Group 1, Group 4, and
Group 5. In the Winter Semester of 2010, there were seventeen students, Group 4, Group 5, and Group 7 had three students, and the other groups had two students. There was one female student in Group 4. In the Fall Semester of 2010, there were thirty-three students while seven were female. Group 3 and Group 5 had four students, and the other groups had five students. There were seven female students who were in Group 3, Group 4, Group 6, and Group 7. In the Winter Semester of 2011, there were forty-two students while seven were female. Each group had six students. Group 4 and Group 6 did not have any female students. For the experiment, Group 3 and Group 5 had seventeen students, and the other groups had eighteen students. Each group had at least one female student.

Table 1

Groups’ Information for Different Semesters

<table>
<thead>
<tr>
<th>Semester</th>
<th>Number of Subjects</th>
<th>Group One</th>
<th>Group Two</th>
<th>Group Three</th>
<th>Group Four</th>
<th>Group Five</th>
<th>Group Six</th>
<th>Group Seven</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2009</td>
<td>32</td>
<td>6 F, 26 M</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Winter 2010</td>
<td>17</td>
<td>1 F, 16 M</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Fall 2010</td>
<td>33</td>
<td>7 F, 26 M</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Winter 2011</td>
<td>42</td>
<td>7 F, 35 M</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>124</td>
<td>21 F, 103 M</td>
<td>18</td>
<td>18</td>
<td>17</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
</tbody>
</table>

Note. F = Female, M = Male. If the group does not have any female, the table does not show any comments for the group.

Table 2 provides a summary of the previous experience of students who participated in the experiment, which may have influenced their Chinese learning.
Table 2

*Previous Experience of Students*

<table>
<thead>
<tr>
<th>Semester</th>
<th>T</th>
<th>HK</th>
<th>S-M</th>
<th>AU</th>
<th>CA</th>
<th>US</th>
<th>NZ</th>
<th>Living Experience</th>
<th>ABC</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall</td>
<td>2009</td>
<td>16</td>
<td>3+1</td>
<td>2</td>
<td>1</td>
<td>---</td>
<td>2</td>
<td>---</td>
<td>5</td>
<td>(2)</td>
</tr>
<tr>
<td>Winter</td>
<td>2010</td>
<td>14</td>
<td>---</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Fall</td>
<td>2010</td>
<td>23</td>
<td>2</td>
<td>---</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>---</td>
<td>2</td>
<td>---</td>
</tr>
<tr>
<td>Winter</td>
<td>2011</td>
<td>24</td>
<td>2</td>
<td>---</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>3+(1)</td>
</tr>
</tbody>
</table>

Total   | 77  | 8  | 3   | 4  | 7  | 4  | 1  | 11                | 3   | 6    |

Note. T = served as full-time Mandarin Chinese-speaking missionaries in Taiwan. HK = served as full-time Cantonese-speaking missionaries in Hong Kong. For fall 2009, the three served their mission in Cantonese, and the one served in Chinese. S-M = served as full-time Mandarin Chinese speaking missionaries in Singapore and Malaysia. AU = served as full-time Mandarin Chinese speaking missionaries in Australia. CA = served as full-time Mandarin Chinese speaking missionaries in Canada. US = served as full-time Mandarin Chinese speaking missionaries in the U.S.A. NZ = served as full-time Mandarin Chinese speaking missionaries in New Zealand. Living Experience = did not serve as full-time Mandarin Chinese-speaking or Cantonese-speaking missionaries and lived in Mainland China at least two months for internship, work, or study, or and studied Chinese several years before coming to BYU. ABC = Chinese heritage who were born in America. For fall 2009, the two in brackets served their Chinese-speaking mission in Taiwan. For winter 2011, the one in brackets served his Chinese-speaking mission in New Zealand. For winter 2011, one of the three that are not in brackets grew up with her adoptive American parents. None = have not served as full-time Mandarin Chinese speaking missionaries and did not live in Mainland China or Taiwan.

The current experiment does not take into account the variable of the former experience of the subjects. It is difficult for us to evaluate the differences of the subjects’ Chinese reading comprehension and the influence of their former experience on their Chinese learning according to their former experience. First, as table 2 showed, there are one hundred four subjects who served as full-time Chinese-speaking or Cantonese-speaking missionaries in Taiwan, Hong Kong, Singapore, Malaysia, Australia, Canada, America, and New Zealand. The situations that they served in as full-time missionaries might have been very different in different countries and areas, although ninety-seven of them were Mandarin Chinese-speaking missionaries. Some of them possibly did not read Chinese every day, while some others possibly studied Chinese characters
by themselves when they communicated with Chinese native speakers. Some of them possibly only used Chinese for the whole time period in which they served their mission, while some of others possibly also spoke other languages. Also, Chinese native speakers that the missionaries contacted during their mission might have effected them quite differently, which may influence their understanding of Chinese language, For example the Chinese native speakers may have had different educational backgrounds, different accents, a different understanding of Chinese history and culture, and so on. Thus, before the students took Chinese 301, in order to achieve the Chinese 301 requirements, some of them only took Chinese 202 at BYU; while some of them took Chinese 201 and 202; and some of them took Chinese 102, 201 and 202. Also, the times between when they served a mission and when they took Chinese 301 are different for these returned missionary students. The shortest was two months, and the longest was more than ten years. This is another reason that it is difficult to evaluate the influence of their Chinese-speaking missions on the reading comprehension test that was held in the Chinese 301 class.

Second, although seventy-seven of them served their mission in Taiwan, the situations in Taiwan were also different. For example, some of them were in Taibei, where many people use Mandarin Chinese, while some of them were in Tainan or Taizhong, where many people prefer to speak a Taiwan dialect that is quite different from Mandarin Chinese.

Third, we cannot definitely say that the students who served their missions using Cantonese in Hong Kong are worse in other areas like Chinese reading comprehension than the students who served their missions using Mandarin Chinese, because in Hong Kong they use simplified Chinese characters.

Fourth, we also cannot definitely say that the students who did not served as full-time Mandarin Chinese-speaking or Cantonese-speaking missionaries are worse than the others. There
are three kinds of students in the table who did not served Mandarin Chinese-speaking missions. One is an American born Chinese, “ABC”. There were two students who grew up with their Chinese-speaking parents, which may have had more influence than full-time Chinese-speaking missions did for language learning. One female student who is of Chinese descent grew up with her adoptive American parents, but from her youth her adoptive American parents tried to let her be involved in the Chinese language and Chinese culture. It is difficult for us to evaluate how much she was influenced by the Chinese language and by Chinese culture before she took Chinese 301.

Another is “Living Experience”. In this group, there was one Korean who was in Chinese elementary school in China before she went to middle school. There was one Singaporean who had studied Chinese for ten years in Singapore and she also went to China before she took Chinese 301 at BYU. There were two Americans who had lived in China with their parents before they took Chinese 301 at BYU. The others lived in Mainland China for either an internship, for working, for studying Chinese, or for family business. It is difficult for us to evaluate the difference between them and the students who served Chinese-speaking missions.

Also, regarding their English background, for the two Koreans and one Singaporean, they were qualified to study in English at American universities, thus I did not separate them from other Americans when they answered the questions in English.

The last one is “None”. These students took Chinese 101, 102, 201, and 202 at BYU before they took Chinese 301. Although they had not served Chinese-speaking missions or had not lived in Mainland China, Taiwan, or other Chinese-speaking areas, they achieved the requirements of Chinese 301 at BYU. Some of them had a good performance in Chinese class.
We cannot say that their former experience did not help them as well as the other students’ former experience.

BYU Chinese 301 is also the first class that does not separate returned missionary students and other students into different sections, which means that returned missionary students and other students are considered to have similar knowledge and abilities within the language, and should be able to meet the same requirements of the course at this level.

There were also six raters who did not know the subjects and who were responsible for evaluating the reading comprehension exercises. As shown in Table 3, all of them were male native American English speakers, and when they evaluated the reading comprehension answers all of them had taken Chinese 302 and were still taking other Chinese classes, which are higher than the Chinese 301 class. They were all either Chinese majors or in the Chinese Flagship program\(^1\) at BYU. Three of them served their two-year mission speaking Mandarin Chinese in Taiwan. One of them served his two-year mission speaking Mandarin Chinese in the area that is predominantly populated by native Chinese speakers in Toronto and had studied Chinese one semester in Mainland China. One of them served his two-year mission speaking Chinese Mandarin in the area that many native Chinese speakers live in New York. Only one rater did not serve his mission using the Chinese language, but he studied Chinese one semester in Mainland China. So, three of them studied Chinese language in mainland China, two of them went to mainland China for other study and travel, and only one rater did not go to mainland China.

\(^1\) The purpose of the Chinese Flagship program is for learners to develop a professional level proficiency in the Chinese language.
Table 3

*Raters’ Chinese Background Information*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>B</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>C</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

Rater A, B, and C evaluated all of the answers of participants in the Fall Semester of 2009 and the Winter Semester of 2010. Rater A evaluated all of the answers of the Fall Semester of 2009 first, and rater B and C evaluated fall 2009 answers respectively with rater A’s rating. Raters A, B, and C evaluated all of the participants’ answer sheets of the Winter Semester of 2010.

Raters D, E, and F evaluated all of the answers of the Fall Semester of 2010 and the Winter Semester of 2011 respectively by using the copies of participants’ answer sheets. Raters E and F were students in Fall Semester of 2009 and participated in the reading comprehension test in 2009, which were evaluated by rater A, B, and C.

We chose these native American English speakers to evaluate the reading answers for the following reasons. First, as their native language is English, they could completely understand participants’ English answers and know if the participants’ English answers corresponded to the questions. Second, after talking with them, the researcher thought their understanding of Chinese was good, and they could understand well the Chinese reading comprehension passages and the key words of the passages at a professional level. Thus, they had good abilities to judge if the participants understood the points that the researcher tested.
Materials

The reading passage that we chose for the experiment is from a Chinese workbook for the third grade of primary school in China. First, we chose some passages from Chinese textbooks and workbooks for Chinese primary school students. We also chose an essay that was written by rater A and was modified by Chinese faculty and a Chinese teaching assistant. The researcher asked some students who were at the beginning of Chinese 302, which is one semester higher than Chinese 301, to read these passages. We asked them to mark the words that they did not know. We also timed them as they read each passage.

According to how many vocabulary words they did not know and the content of the passages, we chose three passages, “màdàjiè kāichē 马大姐开车” (see Appendix C), “zhōngguó kējǔ zhìdù 中国科举制度” (see Appendix G), and “kàn bù jiàn de ài 看不见的爱” (see Appendix K). “màdàjiè kāichē 马大姐开车” and “kàn bù jiàn de ài 看不见的爱” are reading passages in a level-three workbook of a Chinese elementary school. “zhōngguó kējǔ zhìdù 中国科举制度” is written by rater A who is not a native Chinese person. Rater A’s instructor and other teaching assistants modified the passage. According to the words that these Chinese 302 students did not know and what we analyzed that Chinese 301 students may not know, we created a vocabulary list for the three passages. In the list there are Pinyin and English translation of the words.

Rater A, who was taking Chinese 321, made sentences with the vocabulary in the list, and then Chinese TAs modified these sentences. All example sentences were written in both traditional and simplified styles.
We asked rater A to make sentences first, because we think Chinese language students share similar learning experiences. For Chinese 301 students, the sentences that the Chinese 321 student wrote are perhaps more useful than those that would have been written by native speakers.

The reading passage also was written in both traditional and simplified styles. Therefore, the factor of traditional and simplified styles cannot influence the test, since all of the Chinese characters were written in the two styles, and students had access to both.

We chose five words from the list of vocabulary words for each of these three passages, and we used these five words to create five short answer questions in English. We asked English native speakers to modify the questions. Then, we asked some Chinese 302 students to read the passage and answer the questions. According to their answers, we modified the questions again in order to make sure that the questions tested what we wanted to ask.

Students were asked to answer in English, because we are testing Chinese reading comprehension, not Chinese writing. It is possible for some students to understand the passage but not be able to express themselves in Chinese. If we asked them to write in Chinese, we would not know whether the problem had to do with reading comprehension or writing when they made mistakes. The questions were designed in English, because we are only testing students’ reading comprehension of the reading passage, and we did not want their comprehension of the questions to be a variable. It is possible for some students to understand the passage but not understand the questions in Chinese. If the questions were in Chinese, we would not know whether the students did not understand the passage or the questions when they had wrong answers.

In the middle of the Fall Semester of 2009, first, we asked the students to do the reading comprehension exercise of “mǎdàjiě kāichē 马大姐开车” in class and followed the process of
the experiment in order to see if the process worked. “mǎdàijiē kāichē 马大姐开车” has five hundred and nine Chinese characters. The next week, we asked the students to do the reading comprehension exercises of “zhōngguó kējǔ zhìdù 中国科举制度” and also followed the process of the experiment. zhōngguó kējǔ zhìdù 中国科举制度 has six hundred and sixty-one Chinese characters. Then rater A evaluated the answers. The researcher collected and analyzed the data according to the design of analyzing data for the experiment to see if it worked.

At the end of the Fall Semester of 2009, we asked the students to do the reading comprehension exercises of “kàn bù jiàn de ài 看不见的爱”, which the researcher used for analyzing the data. This passage has seven hundred and twenty Chinese characters. Most Chinese 302 students who read the passage spent five to ten minutes reading it, and had less than ten words that they did not know. We created a thirteen-word vocabulary list for the passage. All characters of the list were written in both traditional and simplified styles on the same page.

**Instruments**

As shown in Table 4 below, the current experiment has six treatments combining four different vocabulary learning strategies: contextualized word learning, vocabulary list learning, pre-learning, and normal-learning.

Contextualized word learning in the current experiment consisted of giving students thirteen Chinese words with both English translations and Chinese meaningful sentences that used these target words, because one Chinese word can have different meanings in different situations. For each Chinese target word, the students were given the multiple meanings of the word in translation, including the meaning that was used in the passage. The students were also given a meaningful sentence for the meaning of the target words that are used in the passage.
These meaningful sentences can help students understand the meanings of a certain word in the passage.

Vocabulary list learning in the current experiment consisted of giving students thirteen Chinese words with English translations without meaningful Chinese sentences. The list of target words and English translations is the same one given to the contextualized word learning groups.

Pre-learning vocabulary in the current experiment consists of letting the students study the target words before they read the target passage. In the current study, pre-learning vocabulary can be used with both contextualized word learning and vocabulary list learning.

Normal-learning vocabulary, in this context, means when readers read a text they try to understand the words that they do not know in the text in order to understand the text. In the current experiment consists of letting the students study the target words as they read the target passage. In the current study, normal-learning vocabulary can be used with contextualized word learning, vocabulary list learning, and pre-learning vocabulary.

Table 4

<table>
<thead>
<tr>
<th>Group</th>
<th>Contextualized word learning</th>
<th>Vocabulary list learning</th>
<th>Pre-learning vocabulary</th>
<th>Normal-learning vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Group 2</td>
<td>√</td>
<td>---------</td>
<td>√</td>
<td>---------</td>
</tr>
<tr>
<td>Group 3</td>
<td></td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Group 4</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Group 5</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Group 6</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Group 7</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

As shown in Table 4, Group 2, which is the combination of contextualized word learning and pre-learning vocabulary, was allowed to learn the target words with the English translations
and meaningful sentences before they read the target passage. When they read the target passage, they could not refer back to the vocabulary list or the meaningful sentences.

**Group 3**, which is the combination of vocabulary list learning and pre-learning, was allowed only to learn the target words with English translations before they read the target passage. When they read the target passage, they could not go back to check the list of English translations.

**Group 4**, which is the combination of contextualized word learning, pre-learning, and normal-learning, was allowed to learn the target words with English translations and meaningful sentences before they read the target passage. When they read the target passage, they could refer to the vocabulary list and the meaningful sentences.

**Group 5**, which is the combination of vocabulary list learning, pre-learning, and normal-learning, was asked only to learn the target words with English translations before they read the target passage. When they read the target passage, they could go back to check the list of English translations.

**Group 6**, which is the combination of contextualized word learning and normal-learning, was asked to learn the target words with the English translation and meaningful sentences as they read the target passage.

**Group 7**, which is the combination of vocabulary list learning and normal-learning, was asked to learn only the target words with English translations as they read the target passage.

**Procedures**

Approval to conduct this research was obtained from the Institutional Review Board at Brigham Young University. In order to conduct this study the researcher contacted students in the Chinese 301 language classes at BYU in Fall Semester of 2009, Winter Semester of 2010,
Fall Semester of 2010, and Winter Semester of 2011. During the Fall Semester of 2009, the Fall Semester of 2010, and Winter Semester of 2011, the students attended their normal Chinese classes and did the usual reading comprehension exercises. One of their reading comprehension exercises is the test outlined in this chapter (see Appendix K). At the end of the semester, we briefly explained the study to the students and asked them to sign the Informed Consent Document I (see Appendix A) if they would allow us to analyze their results for this study. Students who allowed us to analyze their results for this study could get an extra ten points in their class.

During the Winter Semester of 2010, the students who wanted to do the extra reading comprehension exercises outlined in this chapter did the test after class. Before they did the test, we asked them to sign the Informed Consent Document II (see Appendix B) in order for us to analyze their results for this study. After they did the test, we briefly explained the study to them, and they were given extra help in their Chinese studies.

In the Fall Semesters of 2009 and 2010 and Winter Semester of 2011, students were randomly arranged into seven groups: Group 1 was the control group; Group 2 read the vocabulary list with example sentences before receiving the reading passage; Group 3 read the vocabulary list before receiving the reading passage; Group 4 read the vocabulary list with example sentences before receiving the reading passage and were allowed to keep it as they read the passage; Group 5 read the vocabulary list before receiving the reading passage and were allowed to keep it as they read the passage; Group 6 read the vocabulary list with example sentences as they read the reading passage; Group 7 read the vocabulary list as they read the reading passage.
The test was held during a normal class period before the final exam as a usual reading comprehension exercise. The students were given different directions (see Appendix O) for the test and were asked to sit according to groups.

First, **Group 2** and **Group 4** were given the vocabulary list with example sentences (see Appendix M). **Group 3** and **Group 5** were given the vocabulary list (see Appendix N). They had five minutes to study their lists. **Group 1**, **Group 6**, and **Group 7** waited quietly for the five minutes.

Second, after the five minutes, the vocabulary lists of **Group 2** and **Group 3** were collected. All groups were given the Chinese reading passage (see Appendix K) and had twelve minutes to read it. **Group 6** was given the vocabulary list with example sentences (see Appendix M) when they were given the Chinese reading passage. **Group 7** was given the vocabulary list (see Appendix N) when they were given the Chinese reading passage. The passage we passed out was written in both traditional Chinese character style and simplified Chinese character style. Students could choose which one they wanted to read.

Third, after twelve minutes, all Chinese passage sheets and vocabulary list sheets were collected. All groups were given questions in English (See Appendix L). They had five minutes to answer these questions in English. After the five minutes, all answer sheets were collected.

In the Winter Semester of 2010, the students who wanted to do extra reading comprehension exercises made appointments and came to the TA office to do the test. When they came, they were randomly arranged into one of the seven groups. They were asked to sign the Informed Consent Document II (see Appendix B) in order to allow us to analyze their results for this study. They were given the instructions of the group into which they had been arranged. After they read the instructions we gave them the test accordingly.
**Scoring**

We found six native English speakers who had a high level of proficiency in Chinese to evaluate the students’ answers, as mentioned above. Before the raters evaluated, they read the passage and answered the questions themselves. We communicated with them about the meaning of the passage and the questions. We also told them the intent of the questions. We explained to them how to grade.

Each question was worth two points. If the students’ answers showed the meaning of the word that we were testing in the passage, they received two points for the question. If the students’ answers showed the meaning of the part that we were testing, but they did not show the exact meaning of the word, they received one point. For example, the answer for the first question is “the mother picked up small stones for her son”. If a student’s answer is that “the mother gave small stones to her son”, which showed that the student understood the context, but he did not write the exact meaning of the word that we were testing, he can receive one point. If the students’ answers did not relate to the question they received zero points.

**Data Analysis**

Each question was evaluated by three raters, and we collected the six raters’ evaluations of the tests that were held during the three semesters. This study used SAS, statistics software, to calculate the mean and standard error for each question and the total scores for the seven groups, and to analyze variance and comparison of different groups according to the two research questions by using the average of the three raters’ evaluation of each question.

**Research Question Number One – Pre-learning Discussion.** The first research question asked: Does pre-learning vocabulary instruction affect CFL (Chinese Foreign Language) learners’ reading comprehension?
As shown in Table 5, the SAS compared Group 2-3 vs. Group 4-5, Group 2-3 vs. Group 6-7, and Group 4-5 vs. Group 6-7 for each question and the total scores, because Group 2-3 represents pre-learning, Group 4-5 represents pre-learning and normal learning, and created Group 6-7 represents normal learning. The SAS also compared the control group and these three created groups respectively, Group 1 vs. Group 2-3, Group 1 vs. Group 4-5, and Group 1 vs. Group 6-7. Thus, we can see not only if there is a significant difference between Group 2-3, Group 4-5, and Group 6-7, but also if there is a significant difference between the control group, Group 1, and the three created treatment groups so that we can analyze if pre-learning has an effect on CFL learners’ reading comprehension.

Table 5

<table>
<thead>
<tr>
<th>Group</th>
<th>Combining</th>
<th>Pre-learning vocabulary</th>
<th>Normal-learning vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2-3</td>
<td>Groups 2 &amp; 3</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Group 4-5</td>
<td>Groups 4 &amp; 5</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Group 6-7</td>
<td>Groups 6 &amp; 7</td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

Research Question Number Two – Contextualized Word Learning Discussion.

Research question two asked: Does teaching contextualized word learning strategy affect CFL learners’ reading comprehension?

As shown in Table 6, the SAS compared combined Group 2-4-6 and combined Group 3-5-7, which represent contextualized word learning and vocabulary list learning respectively. The SAS also compared the control group and the two created treatment groups, Group 1 vs. Group 2-4-6, and Group 1 vs. Group 3-5-7. Thus, we can see not only if there is a significant difference between created Group 2-4-6 and Group 3-5-7, but also if there is a significant
difference between the control group, **Group 1**, and the two created treatment groups so that we can analyze if contextualized word learning has an effect on CFL learners’ reading comprehension. In order to deeply analyze the effect of contextualized word learning, the study also analyzed **Group 2 vs. Group 3, Group 4 vs. Group 5, and Group 6 vs. Group 7** respectively.

Table 6

*Groups for Research Question Two – Contextualized Word Learning*

<table>
<thead>
<tr>
<th>Group</th>
<th>Combining</th>
<th>Contextualized word learning</th>
<th>Vocabulary list learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>---------------</td>
<td>----------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Group 2-4-6</td>
<td>Groups 2, 4, &amp; 6</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Group 3-5-7</td>
<td>Groups 3, 5, &amp; 7</td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

**Other Related Points of Discussion.** In order to analyze the effects of pre-learning and contextualized word learning on CFL learners’ reading comprehension in detail and find possible reasons to explain the results of the two research questions, we also ran SAS to get detailed information about how each group responded to each question. We also compared two groups’ raters, the first-group raters who evaluated the tests of the Fall Semester of 2009 and the Winter Semester of 2010, and the second group raters who evaluated the tests of the Fall Semester of 2010 and the Winter Semester of 2011.

We ran an SAS test to get the distribution of each question and the total scores. Thus, we know the distribution of the readers’ scores for each group and if most students’ scores are close to the mean. Therefore, we have clues of how well the readers were understanding and why they got the wrong answers, which also relates to what the raters were thinking when they evaluated the readers. So, we also compared the means of each rater for each question. We analyzed the correlation of the evaluations of the first three raters and the second three raters so that we can
see the effects of raters on the experiment in order to analyze the experimental results more accurately.
Chapter Four: Results

This study analyzes the two research questions on the effects of pre-learning vocabulary learning and contextualized word learning on reading comprehension. This chapter summarizes and shows the results of data analysis through Statistical Analysis Software (SAS). There are four subsections in this chapter: Basic Information Analysis, Research Question One Analysis, Research Question Two Analysis, and Other Factors Analysis. Basic Information Analysis describes the means of the six treatment groups and one control group for the five questions and the total score. It also shows the \( p \) values of \( t \)-tests of these seven groups for each question and the total score. Research Question One Analysis summarizes the results of the \( t \)-test of the three created groups, Group 2-3, Group 4-5, and Group 6-7, and the control group for each question according to the first research question, which is to analyze the effects of pre-learning vocabulary learning. Research Question Two Analysis summarizes the results of the \( t \)-tests of the two created groups, Group 2-4-6 and Group 3-5-7, and the control group for each question according to the second research question, which is to analyze the effects of contextualized word learning. The other Factors Analysis section shows why we need to analyze the six raters, and it shows the means of each rater’s evaluations and the correlations of these raters with each question.

Basic Data Analysis

The SAS system calculated the mean and standard error of each question and the total scores for the seven groups, which are Group 1: Control; Group 2: Contextualized word learning and pre-learning; Group 3: Vocabulary list learning and pre-learning; Group 4: Contextualized word learning, pre-learning, and normal learning; Group 5: Vocabulary list learning, pre-learning,
and normal learning; Group 6: Contextualized word learning and normal learning; Group 7: Vocabulary list learning and normal learning.

As shown in Table 7, we can see that the mean score for Groups 6 is the highest. The mean scores of Group 1 and Group 7 are the same, which are a little higher than Group 3, which is the lowest. Group 2, Group 4, and Group 7 are close; however, Group 4 is higher than Group 2 and Group 5. From the standard error, we can see that the distributions of the seven groups’ scores are similar, which is relatively large considering the means of the scores. That means that the curve of each group’s distribution is flat, and most students’ scores in each group are not close to the mean.

For question 1, the combination of contextualized word learning and normal-learning that was represented by Group 6 are the best strategies for understanding and answering the first question. The combination of contextualized word learning, pre-learning, and normal learning that was represented by Group 4 is second. The combination of vocabulary list learning and pre-learning vocabulary that was represented by Group 3 is the least helpful strategy, which is even lower than Group 1 that is a no-strategy application. The combination of vocabulary list learning and normal-learning that was represented by Group 7 is the same as Group 1.

Table 7

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.9770</td>
<td>1.1437</td>
<td>0.9474</td>
<td>1.1992</td>
<td>1.1631</td>
<td>1.3289</td>
<td>0.9770</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.1840</td>
<td>0.1840</td>
<td>0.1891</td>
<td>0.1840</td>
<td>0.1891</td>
<td>0.1840</td>
<td>0.1840</td>
</tr>
</tbody>
</table>

Table 8 provides a summary of p value of t-tests for question 1. As they are all much higher than 0.05 and most of them are close to 1 or are 1, this means that there is no significant difference between each of the two groups. Thus, the best groups and the worst groups are not
significantly different, and there is no one reading strategy that is significantly different from another reading strategy or no-strategy application as shown in Group 1 for question 1.

Table 8

*P value of t-test for Question 1*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.9952</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.0000</td>
<td>0.9893</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.9780</td>
<td>1.0000</td>
<td>0.9620</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.9919</td>
<td>1.0000</td>
<td>0.9837</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.8231</td>
<td>0.9915</td>
<td>0.7733</td>
<td>0.9988</td>
<td>0.9957</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.0000</td>
<td>0.9952</td>
<td>1.0000</td>
<td>0.9780</td>
<td>0.9919</td>
<td>0.8231</td>
</tr>
</tbody>
</table>

As shown in Table 9, we can see the mean of Group 2’s scores is the highest, and Group 5 is the second highest. Group 3, Group 4, and Group 6 are close, which is a little lower than Group 7. The mean of Group 1’s scores is the lowest. From the standard error, we see the distributions of the seven groups’ scores are similar, which is relatively large considering the mean of the scores. This means that the curve of each group’s distribution is flat, and most students’ scores in each group are not close to the mean.

For question 2, the combination of contextualized word learning and pre-learning that was represented by Group 2 is the best method. The combination of vocabulary list learning, pre-learning, and normal learning is in the second place, which is also good for answering question 2. All of the treatments are better than no strategy application for question 2.

Table 9

*Means of Question 2*

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.7374</td>
<td>1.4411</td>
<td>1.0415</td>
<td>1.0522</td>
<td>1.2768</td>
<td>1.0708</td>
<td>1.1633</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.1985</td>
<td>0.1985</td>
<td>0.2040</td>
<td>0.1985</td>
<td>0.2040</td>
<td>0.1985</td>
<td>0.1985</td>
</tr>
</tbody>
</table>
As shown in Table 10, we see that most $p$ values are much higher than 0.05, and some of them are close to or are 1, which means that there is no significant difference between each of the groups compared, and some of the groups are very similar, such as Group 3 and Group 4, Group 3 and Group 6, and Group 4 and Group 6. Although the mean of Group 2, which is the highest, is much higher than the mean of Group 1, which is lowest, the $p$ value of $t$-test for Group 1 and Group 2 shows that it is still not a significant difference. Thus, the combination of contextualized word learning and pre-learning is not significantly different from no-strategy and other strategies application for question 2.

Table 10

$P$ value of $t$-test for Question 2

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.1632</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.9352</td>
<td>0.7964</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.9190</td>
<td>0.8059</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.4842</td>
<td>0.9973</td>
<td>0.9827</td>
<td>0.9854</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.8960</td>
<td>0.8393</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9907</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.7306</td>
<td>0.9546</td>
<td>0.9995</td>
<td>0.9997</td>
<td>0.9997</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

As shown in Table 11, we see that the mean of Group 4’s scores is the highest, which is much higher than the mean of Group 6, which is the lowest. Group 6 is even a little lower than Group 1. The mean of Group 7 is the second highest. The means of Groups 3 and 5 are very close and are lower than Group 7. From the standard error, we see the distributions of the seven groups’ scores are similar, which is relatively large considering the mean of the scores. This means that the curve of each group’s distribution is flat, and most students’ scores in each group are not close to the mean.
The combination of contextualized word learning, pre-learning, and normal learning is the best strategy for understanding and answering question 3. The combination of contextualized word learning and normal-learning is the worst strategy to apply to question 3.

Table 11

*Means of Question 3*

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.7768</td>
<td>0.9250</td>
<td>1.0434</td>
<td>1.4065</td>
<td>1.0238</td>
<td>0.7398</td>
<td>1.2213</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.1408</td>
<td>0.1408</td>
<td>0.1446</td>
<td>0.1408</td>
<td>0.1446</td>
<td>0.1408</td>
<td>0.1408</td>
</tr>
</tbody>
</table>

As shown in Table 12, we can see most p values are much higher than 0.05, some of them are close to or are 1, which means that there is no significant difference between each of the two groups compared, and some of the groups are very similar, such as Group 1 and Group 6. However, the mean of Group 4, which is the highest, is much higher than the mean of Group 6, which is the lowest, and the p value, 0.0177, is less than 0.05, which means that it is significantly different. The p value of Group 1 and Group 4 is 0.0308, which is also significantly different. The p values of Group 2 and Group 4, Group 6 and Group 7, and Group 1 and Group 7 are relatively low, but they are still not significantly different. Thus, for question 3, the combination of contextualized word learning, pre-learning, and normal learning is the strategy that is significantly different from the combination of the contextualized word learning and normal learning and no-strategy application that is represented by Group 1.
Table 12

*P value of t-test for Question 3*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.9892</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.8392</td>
<td>0.9970</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.0308</td>
<td>0.1968</td>
<td>0.5481</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.8823</td>
<td>0.9989</td>
<td>1.0000</td>
<td>0.4835</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.0000</td>
<td>0.9663</td>
<td>0.7393</td>
<td>0.0177</td>
<td>0.7950</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.2823</td>
<td>0.7480</td>
<td>0.9743</td>
<td>0.9663</td>
<td>0.9571</td>
<td>0.1968</td>
</tr>
</tbody>
</table>

As shown in Table 13, we can see the mean of Group 4’s scores is the highest, which is much higher than the mean of Group 2, which is the lowest. Group 2 is even worse than Group 1. Group 5 and Group 6 are close, which are lower than the Group 4. Group 3 and Group 7 are close, which are higher than Group 1. From the standard error, we can see the distributions of the seven groups’ scores are similar, which is relatively large considering the mean of the scores. That means that the curve of each group’s distribution is flat, and most students’ scores in each group are not close to the mean.

The combination of contextualized word learning, pre-learning, and normal learning is the best strategy to answer question 4. The combination of contextualized word learning and pre-learning vocabulary is the worst one, and it is worse than no strategy for question 4.

Table 13

*Means of Question 4*

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.7736</td>
<td>0.6440</td>
<td>0.8350</td>
<td>1.1440</td>
<td>0.9918</td>
<td>0.9588</td>
<td>0.8662</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.1415</td>
<td>0.1415</td>
<td>0.1454</td>
<td>0.1415</td>
<td>0.1454</td>
<td>0.1415</td>
<td>0.1415</td>
</tr>
</tbody>
</table>

As shown in Table 14, we see that most *p* values are much higher than 0.05, and some of them are close to 1 while the *p* value of Group 3 and Group 7 is 1. This means that there is no significant difference between each of the two groups, and some of the groups are very similar.
The mean of Group 4, which is the highest, is much higher than the mean of Group 2, which is lowest, and the \( p \) value, 0.1659, is higher than 0.05, which means that it is still not significantly different. Therefore, the combination of contextualized word learning, pre-learning, and normal learning can be considered a good strategy for question 4.

Table 14

\[ P \text{ value of } t\text{-test for Question 4} \]

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.9949</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.9999</td>
<td>0.9644</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.5121</td>
<td>0.1659</td>
<td>0.7274</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.9332</td>
<td>0.6044</td>
<td>0.9878</td>
<td>0.9888</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.9671</td>
<td>0.6957</td>
<td>0.9963</td>
<td>0.9671</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.9992</td>
<td>0.9224</td>
<td>1.0000</td>
<td>0.8042</td>
<td>0.9960</td>
<td>0.9992</td>
</tr>
</tbody>
</table>

As shown in Table 15, we can see the mean of Group 4’s scores is the highest, which is much higher than the mean of Group 2, which is the lowest. The means of other groups are similar. From the standard error, we see that the distribution of the seven groups’ scores are similar, which is relatively large considering the mean of the scores. That means that the curve of each group’s distribution is flat, and most students’ scores in each group are not close to the mean.

The combination of contextualized word learning, pre-learning, and normal learning is the best strategy for understanding and answering question 5. The combination of contextualized word learning and pre-learning vocabulary is the worst strategy for question 5; it is even worse than the no strategy application.
Table 15

Means of Question 5

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.2099</td>
<td>0.9136</td>
<td>1.3187</td>
<td>1.4692</td>
<td>1.3383</td>
<td>1.3581</td>
<td>1.3395</td>
</tr>
<tr>
<td></td>
<td>0.1552</td>
<td>0.1552</td>
<td>0.1594</td>
<td>0.1552</td>
<td>0.1594</td>
<td>0.1552</td>
<td>0.1552</td>
</tr>
</tbody>
</table>

As shown in Table 16, we see that most $p$ values are much higher than 0.05, some of them are close to or are 1, which means that there is no significant difference between each of the groups compared, and some of the groups are very similar, such as Group 3 and Group 5, Group 3 and Group 6, Group 3 and Group 7, Group 5 and Group 6, Group 5 and Group 7, and Group 6 and Group 7. The mean of Group 4, which is the highest, is higher than the mean of Group 2, which is the lowest, and the $p$ value, 0.1547, is higher than 0.05, which means that it is still not significantly different. Thus, the combination of contextualized word learning, pre-learning, and normal learning is not significantly different from no-strategy and other strategies application for question 5.

Table 16

$P$ value of $t$-test for Question 5

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.8241</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.9989</td>
<td>0.5337</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.8982</td>
<td>0.1547</td>
<td>0.9935</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.9973</td>
<td>0.4754</td>
<td>1.0000</td>
<td>0.9970</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.9936</td>
<td>0.4001</td>
<td>1.0000</td>
<td>0.9987</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.9969</td>
<td>0.4534</td>
<td>1.0000</td>
<td>0.9969</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

As shown in Table 17, for the total score, we see that the mean of Group 4’s scores, 6.2711, is the highest, and the mean of Group 1’s scores is the lowest. The mean of Group 5, 5.7937, is the second highest, which is a little higher than the mean of Group 7 and Group 6, which are the third and the fourth highest. The means of Group 3 and Group 2, which are the
second and the third lowest, are higher than the mean of Group 1, 4.4748. From the standard error, we see that the distribution of the seven groups’ scores are similar, which is relatively small considering the mean of the scores. This means that the curve of each group’s distribution is relatively high, and most students’ scores in each group are close to the mean.

All strategies are better than no strategy application for understanding the whole passage and answering the five reading comprehension questions, while the combination of contextualized word learning, pre-learning, and normal learning is better than the other strategies.

Table 17

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>4.4748</td>
<td>5.0674</td>
<td>5.1859</td>
<td>6.2711</td>
<td>5.7937</td>
<td>5.4563</td>
<td>5.5674</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.4994</td>
<td>0.4994</td>
<td>0.5131</td>
<td>0.4994</td>
<td>0.5131</td>
<td>0.4994</td>
<td>0.4994</td>
</tr>
</tbody>
</table>

As shown in Table 18, for the total score, we see that most p values are much higher than 0.05, some of them are close to or are 1, which means that there is no significant difference between each of the groups compared, and some of the groups are very similar, such as Group 2 and Group 3, and Group 6 and Group 7. Although the mean of Group 4, which is the highest, is higher than the mean of Group 1, which is the lowest, the p value, 0.1505, is higher than 0.05, which means that it is still not significantly different. The p values of Group 1 and Group 5 and Group 2 and Group 4 are relatively low, but they are still not significantly different.

The combination of contextualized word learning, pre-learning, and normal learning is better than other combinations and the no strategy application, but the difference is not statistically significant.
Table 18

*P value of t-test for Total Score*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.9799</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.9539</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.1505</td>
<td>0.6102</td>
<td>0.7322</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.5191</td>
<td>0.9491</td>
<td>0.9802</td>
<td>0.9940</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.8035</td>
<td>0.9979</td>
<td>0.9998</td>
<td>0.9081</td>
<td>0.9991</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.7122</td>
<td>0.9917</td>
<td>0.9983</td>
<td>0.9530</td>
<td>0.9999</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Table 19 provides a summary of groups of highest and lowest means for each question and total score.

Table 19

*Groups of Highest and Lowest Means for Each Question and Total Score*

<table>
<thead>
<tr>
<th>Question</th>
<th>Groups of the Highest Mean</th>
<th>Groups of the Second Highest Mean</th>
<th>Groups of the Lowest Mean</th>
<th>Groups of the Second Lowest Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group</td>
<td>Mean</td>
<td>Group</td>
<td>Mean</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>1.3289</td>
<td>4</td>
<td>1.1992</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1.4411</td>
<td>5</td>
<td>1.2768</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>1.4065</td>
<td>7</td>
<td>1.2213</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1.1440</td>
<td>5</td>
<td>0.9918</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1.4692</td>
<td>6</td>
<td>1.3581</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>6.2711</td>
<td>5</td>
<td>5.7937</td>
</tr>
</tbody>
</table>

Although the order of the means for different questions is not consistent, we still can see Group 4 had the best performance. Group 4 is in the first place three times for the five questions and has the highest mean for the total score. Group 5’s performance can be considered as good too, as Group 5 is three times in the second place. Therefore, compared with other strategies, the combination of contextualized word learning, pre-learning, and normal learning is the best strategy, and the combination of vocabulary list learning, pre-learning, and normal-learning is also a good strategy.
Although the mean of Group 1 for the total score is the lowest, we cannot definitely say that only Group 1 has the worst performance because of the following reasons. First, Group 2 is also in last place twice. Second, Group 3 is even worse than Group 1 for the first question. The mean of the total score is close to the mean of Group 2, which is the last second place. Third, most $p$ values of the $t$-test of Group 1 & Group 2, Group 1 & Group 3, and Group 2 & Group 3, are 1 or are close to 1 for the five questions and the total score, which means they are very similar to each other. Only the $p$ value of the $t$-test of Group 1 & Group 2 for question 2 is low, 0.1632, which means the difference between Group 1 and Group 2 for question 2 is significant. For question 5, although the $p$ value of the $t$-test of Group 2 & Group 3 is much lower than other $p$ values, it is still not significant. Also, the $p$ values of $t$-test of Group 1 & Group 2 and Group 1 & Group 3 for question 5 are very high, which means they are very close to each other. Thus, the combination of contextualized word learning and pre-learning vocabulary and the combination of vocabulary list learning and pre-learning also do not have good performances.

Although only the $p$ values of question 3 show that there are significant differences between Group 4 & Group 1 and Group 4 & Group 6, from Table 14, we see that the means of Group 4 are higher than other groups, especially the lowest groups. From $p$ values of Group 4 & Group 2 for question 4 (0.1659), Group 4 & Group 2 for question 5 (0.1547), Group 4 & Group 1 for the total score (0.1505), we can see that Group 4 is different from other groups, although the differences are not statistically significant. The $p$ values of Group 5 & Group 1 for question 2 and Group 5 & Group 2 for question 5 are relatively low, which show that Group 5 is also different from the last group, although the differences are not significant.

In short, we can say that from these tables Group 4, which is the combination of contextualized word learning, pre-learning, and normal learning, is the most effective group.
Group 5, which combined vocabulary list learning, pre-learning, and normal-learning, is also good and is just below Group 4. Group 6, which combined contextualized word learning and normal-learning, and Group 7, which combined vocabulary list learning & normal-learning, are similar, and these two groups are between the first place and the last place. The means of Group 1, which is the control group, Group 2, which is the combination of contextualized word learning and pre-learning vocabulary, and Group 3, which is the combination of vocabulary list learning and pre-learning, are close and they are the last three groups.

**Research Question One Analysis**

The first research question asked: *Does pre-learning vocabulary instruction affect CFL (Chinese Foreign Language) learners’ reading comprehension?*

In order to analyze the effect of pre-learning, we created **Group 2-3** by combining Group 2 and Group 3, to represent pre-learning; **Group 4-5** by combining Group 4 and Group 5 to represent pre-learning and normal-learning; and **Group 6-7** by combining Group 6 and Group 7 to represent normal-learning. We used SAS and did a t-test for these three created groups and Group 1, the control group.

As shown in Table 20, for question 1, we can see that the means of created Group 2-3, Group 4-5, and Group 6-7 are higher than the mean of Group 1. The mean of Group 2-3 is lower than the mean of Group 4-5 and Group 6-7, while the mean of Group 4-5 is higher than the mean of Group 6-7. However, all the *p* values are much higher than 0.05, which shows that there is no significant difference between each of the two groups.

The combination of pre-learning and normal-learning, which is represented by created Group 4-5, is better than other strategies, while no strategy application is the worst for the first
question. However, there is no significant difference from the combination of pre-learning and normal-learning to the other strategies and to no strategy.

Table 20

*T-test of Question 1 for Analyzing Pre-Learning*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs. 2-3</td>
<td>-0.06854375</td>
<td>0.22562982</td>
<td>-0.30</td>
<td>0.7618</td>
<td></td>
</tr>
<tr>
<td>1 vs. 4-5</td>
<td>-0.20416467</td>
<td>0.22562982</td>
<td>-0.90</td>
<td>0.3674</td>
<td></td>
</tr>
<tr>
<td>1 vs. 6-7</td>
<td>-0.17592593</td>
<td>0.22452597</td>
<td>-0.78</td>
<td>0.4349</td>
<td></td>
</tr>
<tr>
<td>2-3 vs. 4-5</td>
<td>-0.13562092</td>
<td>0.18600110</td>
<td>-0.73</td>
<td>0.4674</td>
<td></td>
</tr>
<tr>
<td>2-3 vs. 6-7</td>
<td>-0.10738217</td>
<td>0.18467498</td>
<td>-0.58</td>
<td>0.5621</td>
<td></td>
</tr>
<tr>
<td>4-5 vs. 6-7</td>
<td>0.02823874</td>
<td>0.18467498</td>
<td>0.15</td>
<td>0.8787</td>
<td></td>
</tr>
</tbody>
</table>

As shown in Table 21, for question 2, we see that the means of created Group 2-3, Group 4-5, and Group 6-7 are higher than the mean of Group 1. The mean of Group 2-3 is higher than the mean of Group 4-5 and Group 6-7, while the mean of Group 4-5 is higher than the mean of Group 6-7. The *p* value of Group 1 vs. Group 2-3 shows that there is a significant difference between these two groups. The *p* value of Group 1 vs. Group 4-5 (0.08) is a little higher than 0.05, which shows that there is a difference which approaches significance. All other *p* values are higher than 0.05, which means that there is not a significant difference between the other group comparisons.

Pre-learning, which is represented by Group 2-3, is better than other strategies, while the no strategy application is the worst for the second question. Pre-learning is significantly different from the no strategy application.
Table 21

_T-test of Question 2 for Analyzing Pre-Learning_

| Parameter | Estimate          | Standard Error | t Value | Pr > |t| |
|-----------|-------------------|----------------|---------|------|---|
| 1 vs. 2-3 | -0.50388203       | 0.24340541     | -2.07   | 0.0407 |
| 1 vs. 4-5 | -0.42708465       | 0.24340541     | -1.75   | 0.0820 |
| 1 vs. 6-7 | -0.37962963       | 0.24221460     | -1.57   | 0.1198 |
| 2-3 vs. 4-5 | 0.07679739     | 0.20065466     | 0.38    | 0.7026 |
| 2-3 vs. 6-7 | 0.12425240     | 0.19922406     | 0.62    | 0.5341 |
| 4-5 vs. 6-7 | 0.04745502     | 0.19922406     | 0.24    | 0.8121 |

As shown in Table 22, for question 3, we can see that the means of created Group 2-3, Group 4-5, and Group 6-7 are higher than the mean of Group 1. The mean of Group 2-3 is lower than the mean of Group 4-5, but it is a little higher than Group 6-7. The \( p \) value of Group 1 vs. Group 4-5 is 0.0124 that shows there is a significant difference between these two groups. The \( p \) values of Group 4-5 vs. Group 6-7 and Group 2-3 vs. Group 4-5 are relatively low, but they are not significant. All other \( p \) values are higher than 0.05, which means that there is not a significant difference between each of the two groups. In particular, the \( p \) value of Group 2-3 vs. Group 6-7 is very close to 1, which means that the means of Group 2-3 and Group 6-7 are similar.

The combination of pre-learning and normal-learning, which is represented by created Group 4-5, is better than other strategies, while the no strategy application is the worst for the third question. There is a significant difference between the combination of pre-learning and normal-learning and no strategy. There are also differences from the combination of pre-learning and normal-learning to the other strategies, but the differences are not statistically significant.
Table 22

*T-test of Question 3 for Analyzing Pre-Learning*

| Parameter | Estimate | Standard Error | t Value | Pr > |t| |
|-----------|----------|----------------|---------|------|---|
| 1 vs. 2-3 | -0.20734011 | 0.17262313 | -1.20 | 0.2322 |
| 1 vs. 4-5 | -0.43827693 | 0.17262313 | -2.54 | 0.0124 |
| 1 vs. 6-7 | -0.20370370 | 0.17177860 | -1.19 | 0.2381 |
| 2-3 vs. 4-5 | -0.23093682 | 0.14230429 | -1.62 | 0.1073 |
| 2-3 vs. 6-7 | 0.00363641 | 0.14128971 | 0.03 | 0.9795 |
| 4-5 vs. 6-7 | 0.23457323 | 0.14128971 | 1.66 | 0.0996 |

As shown in Table 23, for question 4, we can see that the mean of Group 1 is lower than the means of created Group 4-5 and Group 6-7, but it is higher than the mean of Group 2-3. The mean of Group 4-5 is higher than Group 6-7. Most p values are higher than 0.05, which means that there is no significant difference between each of the two groups. However, the p value of Group 2-3 vs. Group 4-5 is 0.0234, which is statistically significant. The p value of Group 1 vs. Group 4-5 is 0.0925, which is a little higher than 0.05, and shows that there is a difference, but it indicates no significant difference.

The combination of pre-learning and normal-learning, which is represented by created Group 4-5, is better than other strategies, and pre-learning, which is represented by created Group 2-3, is worse than other strategies and no-strategy for question 4. The combination of pre-learning and normal-learning is significantly different from pre-learning. The no strategy application is a little better than pre-learning; however, it has a large difference from the combination of pre-learning and normal-learning.
Table 23

*T*-test of Question 4 for Analyzing Pre-Learning

| Parameter        | Estimate       | Standard Error | t Value | Pr > |t| |
|------------------|----------------|----------------|---------|------|---|
| 1 vs. 2-3        | 0.03416063     | 0.17347171     | 0.20    | 0.8442 |
| 1 vs. 4-5        | -0.29427074    | 0.17347171     | -1.70   | 0.0925 |
| 1 vs. 6-7        | -0.13888889    | 0.17262303     | -0.80   | 0.4227 |
| 2-3 vs. 4-5      | -0.32843137    | 0.14300383     | -2.30   | 0.0234 |
| 2-3 vs. 6-7      | -0.17304952    | 0.14198426     | -1.22   | 0.2254 |
| 4-5 vs. 6-7      | 0.15538185     | 0.14198426     | 1.09    | 0.2761 |

As shown in Table 24, for question 5, we can see that the mean of Group 1 is lower than the means of created Group 4-5 and Group 6-7, but it is higher than the mean of Group 2-3. The mean of Group 4-5 is higher than Group 6-7. Most p values are much higher than 0.05, which means that there is not a significant different between each of the two groups. However, the p value of Group 2-3 vs. Group 4-5 is 0.0693, which is close to 0.05. The p value shows that there is a difference between Group 4-5 and Group 2-3, and the difference approaches significance. The p value of Group 2-3 vs. Group 6-7 is relatively low, but it is not significant.

The combination of pre-learning and normal-learning, which is represented by created Group 4-5, is better than other strategies, and pre-learning, which was represented by created Group 2-3, is worse than other strategies for question 4, and it is even worse than the no strategy application. The combination of pre-learning and normal-learning is different from pre-learning alone, but it is not statistically significant. Normal learning, which was represented by created Group 6-7, is in second place for the last question, and it is not significantly different from pre-learning.
Table 24

*T-test of Question 5 for Analyzing Pre-Learning*

| Parameter   | Estimate  | Standard Error  | t Value | Pr > |t| |
|-------------|-----------|-----------------|---------|------|---|
| 1 vs. 2-3   | 0.09377655| 0.19029102      | 0.49    | 0.6231|
| 1 vs. 4-5   | -0.19380515| 0.19029102      | -1.02   | 0.3106|
| 1 vs. 6-7   | -0.13888889| 0.18936005      | -0.73   | 0.4648|
| 2-3 vs. 4-5 | -0.28758170| 0.15686906      | -1.83   | 0.0693|
| 2-3 vs. 6-7 | -0.23266544| 0.15575064      | -1.49   | 0.1379|
| 4-5 vs. 6-7 | 0.05491626 | 0.15575064      | 0.35    | 0.7250|

As shown in Table 25, for the total score, we can see that the mean of Group 1 is lower than the means of created Group 2-3, Group 4-5, and Group 6-7. The mean of Group 2-3 is lower than the mean of Group 4-5 and Group 6-7, while Group 4-5 is higher than Group 6-7. The *p* values of Group 2-3 vs. Group 6-7, Group 4-5 vs. Group 6-7, and Group 1 vs. Group 2-3, are higher than 0.05, which means that it there is not a significant different between each of the groups. However, the *p* value of Group 1 vs. Group 4-5 is 0.0123, which is lower than 0.05. The *p* value shows that the difference between Group 1 and Group 4-5 is significantly different. The *p* value of Group 2-3 vs. Group 4-5 is 0.0753, which is a little higher than 0.05, and shows that there is a difference between Group 2-3 and Group 4-5, and the difference approaches significance.

For the total score, the combination of pre-learning and normal-learning, which is represented by created Group 4-5, is better than other strategies, and the no strategy application, which was represented by Group 1, is worse than other strategies. The combination of pre-learning and normal-learning is significantly different from the no strategy application. The combination of pre-learning and normal-learning is better than pre-learning, but it is still not a significant difference.
Table 25

*T-test of Total Score for Analyzing Pre-Learning*

| Parameter       | Estimate       | Standard Error | t Value | Pr > |t| |
|-----------------|----------------|----------------|---------|-------|---|
| 1 vs. 2-3       | -0.65182872    | 0.61229491     | -1.06   | 0.2893|   |
| 1 vs. 4-5       | -1.55760214    | 0.61229491     | -2.54   | 0.0123|   |
| 1 vs. 6-7       | -1.03703704    | 0.60929936     | -1.70   | 0.0914|   |
| 2-3 vs. 4-5     | -0.90577342    | 0.50475387     | -1.79   | 0.0753|   |
| 2-3 vs. 6-7     | -0.38520832    | 0.50115515     | -0.77   | 0.4437|   |
| 4-5 vs. 6-7     | 0.52056510     | 0.50115515     | 1.04    | 0.3011|   |

Table 26 provides a summary of groups of highest and lowest means for each question and total score for pre-learning.

Table 26

*Groups of Highest and Lowest Means for Each Question and Total Score for Pre-Learning*

<table>
<thead>
<tr>
<th>Questions</th>
<th>Groups of the Highest Mean</th>
<th>Groups of the lowest Mean</th>
<th>Estimate</th>
<th>P Value of t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Group 4-5</td>
<td>Group 1</td>
<td>0.20416467</td>
<td>0.3674</td>
</tr>
<tr>
<td>2</td>
<td>Group 2-3</td>
<td>Group 1</td>
<td>0.50388203</td>
<td>0.0407</td>
</tr>
<tr>
<td>3</td>
<td>Group 4-5</td>
<td>Group 1</td>
<td>0.43827693</td>
<td>0.0124</td>
</tr>
<tr>
<td>4</td>
<td>Group 4-5</td>
<td>Group 2-3</td>
<td>0.32843137</td>
<td>0.0234</td>
</tr>
<tr>
<td>5</td>
<td>Group 4-5</td>
<td>Group 2-3</td>
<td>0.28758170</td>
<td>0.0693</td>
</tr>
<tr>
<td>Total</td>
<td>Group 4-5</td>
<td>Group 1</td>
<td>1.55760214</td>
<td>0.0123</td>
</tr>
</tbody>
</table>

The order of the means for different questions and the total score is relatively consistent, and we can see that Group 4-5 is better than the other three groups. Group 4-5 is in first place four times for the five questions and has the highest mean for the total score. The differences of means between Group 4-5 and the group that is in the last place have a difference range from 0.2042 to 0.5039 for the four questions, with two points being the full points for each question. Group 4-5 is 1.5576 higher than Group 1, which is in the last place for the total score, with ten points being the full points. There are three times that the difference between Group 4-5 and the
group that is in the last place is significant. Therefore, the combination of pre-learning and normal-learning, which is represented by created Group 4-5, is the best strategy in the test.

Group 1 is in last place four times for the three questions and the total score, and there are three times that the difference between its mean and the mean of the group that is in first place, is significant. Group 1 is 0.5039 lower than Group 4-5 for question 2, of which the two points are the full points. Group 2-3 is in the last place twice, and both times are significantly different from Group 4-5. However, Group 2-3 is in the first place for question 2, which is a significant difference from Group 1. Therefore, any strategy is better than no strategy application, although pre-learning does not show a very strong positive effect.

Table 27 provides a summary of significantly different groups for each question and total score for pre-learning. We can see that for question 3, question 4, and the total score, not only do Group 4-5 and the group that is in the last place have a significant difference, but also other groups are significantly different from each other. For question 3, Group 4-5 has a difference from not only Group 1 but also Group 2-3 and Group 6-7, although the differences from Group 2-3 and Group 6-7 are not statistically significant. For question 4, although Group 1 is a little better than the last place—Group 2-3—it still has a significant difference from Group 4-5. For the total score, Group 2-3 is the second to last place, and it still has a difference from the first place—Group 4-5, although the difference cannot be considered as statistically significant.

For question 5, the difference between Group 4-5 and Group 2-3, which is the last group, approaches significance ($p = .0693$). For question 2, the mean of Group 4-5 is a little lower than the mean of Group 2-3, but Group 4-5 is still different from Group 1 that is in the last place, although the difference is not statistically significant.
Therefore, the effect of pre-learning combined with normal learning, which was represented by Group 4-5, is significant and the strategy is effective.

Table 27

*Significantly Different Groups for Each Question and Total Score for Pre-Learning*

<table>
<thead>
<tr>
<th>Questions</th>
<th>Groups</th>
<th>Estimate</th>
<th>( P ) Value of ( t )-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-------</td>
<td>---------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>2</td>
<td>Group 1 vs. Group 2-3</td>
<td>-0.50388203</td>
<td>0.0407</td>
</tr>
<tr>
<td></td>
<td>(Group 1 vs. Group 4-5)</td>
<td>(-0.42708465)</td>
<td>(0.0820)</td>
</tr>
<tr>
<td>3</td>
<td>Group 1 vs. Group 4-5</td>
<td>-0.43827693</td>
<td>0.0124</td>
</tr>
<tr>
<td></td>
<td>(Group 4-5 vs. Group 6-7)</td>
<td>(0.23457323)</td>
<td>(0.0996)</td>
</tr>
<tr>
<td></td>
<td>(Group 2-3 vs. Group 4-5)</td>
<td>(-0.23093682)</td>
<td>(0.1073)</td>
</tr>
<tr>
<td>4</td>
<td>Group 2-3 vs. Group 4-5</td>
<td>-0.32843137</td>
<td>0.0234</td>
</tr>
<tr>
<td></td>
<td>(Group 1 vs. Group 4-5)</td>
<td>(-0.29427074)</td>
<td>(0.0925)</td>
</tr>
<tr>
<td>5</td>
<td>(Group 2-3 vs. Group 4-5)</td>
<td>(-0.28758170)</td>
<td>(0.0693)</td>
</tr>
<tr>
<td>Total</td>
<td>Group 1 vs. Group 4-5</td>
<td>-1.55760214</td>
<td>0.0123</td>
</tr>
<tr>
<td></td>
<td>(Group 2-3 vs. Group 4-5)</td>
<td>(-0.90577342)</td>
<td>(0.0753)</td>
</tr>
<tr>
<td></td>
<td>(Group 1 vs. Group 6-7)</td>
<td>(-1.03703704)</td>
<td>(0.0914)</td>
</tr>
</tbody>
</table>

Note. The \( p \) values that are more than 0.05 but lower than 0.1 were listed in brackets.

Although for question 2, Group 2-3 is in the first place and has a significant difference from Group 1 that is the last group, Group 2-3 does not show a strong positive effect. Group 2-3 is in the last place for question 4 and question 5 and significantly different from Group 4-5 that is in the first place. For question 3, Group 2-3 is very close to Group 6-7, and both of them have a difference from Group 4-5 that is in the first group, although the difference cannot be considered as statistically significant. For the total score, the \( p \) value of Group 2-3 vs. Group 4-5 is 0.0753, which is a little higher 0.05 showing that there is a difference between these two groups, and the
difference approaches significance. Therefore, the effect of pre-learning, which is represented by Group 2-3, is very small, and pre-learning is almost the same as the no strategy application.

**Research Question Two Analysis**

The second research question asked: *Does teaching contextualized word learning strategy affect CFL (Chinese Foreign Language) learners’ reading comprehension?*

In order to analyze the effect of contextualized word learning, we created Group 2-4-6 by combining Group 2, Group 4, and Group 6 to represent contextualized word learning. We also created Group 3-5-7 by combining Group 3, Group 5, and Group 7 to represent vocabulary list learning. We used SAS and did a t-test for each of these two created groups and Group 1, the control group.

As shown in Table 28, for question 1, we can see that the means of created Group 2-4-6 and Group 3-5-7 are higher than the mean of Group 1. The mean of Group 3-5-7 is lower than the mean of Group 2-4-6. However, although the $p$ values of Group 2-4-6 vs. Group 3-5-7 and Group 1 vs. Group 2-4-6 are relatively low, all the $p$ values are much higher than 0.05, which shows that there is no significant difference between the groups.

The contextualized word learning that is represented by created Group 2-4-6 is better than the vocabulary list learning represented by created Group 3-5-7, while the no strategy application is the worst for the first question. However, there is no significant difference between them.
Table 28

*T-test of Question 1 for Analyzing the Contextualized Word Learning*

| Parameter       | Estimate  | Standard Error | t Value | Pr > |t| |
|-----------------|-----------|----------------|---------|------|---|
| 2-4-6 vs. 3-5-7 | 0.19473565| 0.15115854     | 1.29    | 0.2002 |
| 1 vs. 2-4-6     | -0.2469111| 0.21168458     | -1.17   | 0.2458 |
| 1 vs. 3-5-7     | -0.05217546| 0.21273140    | -0.25   | 0.8067 |

As shown in Table 29, for question 2, we can see that the means of created Group 2-4-6 and Group 3-5-7 are higher than the mean of Group 1. The mean of Group 3-5-7 is a little lower than the mean of Group 2-4-6. The p values of Group 1 vs. Group 2-4-6 (0.0508), which is very close to 0.05, shows that the difference between the groups approaches significance. The p value of Group 1 vs. Group 3-5-7 is 0.0678, which is close to 0.05, and shows that the difference between Group 1 and Group 3-5-7 approaches significance. The p value of Group 2-4-6 vs. Group 3-5-7 is much higher than 0.05, which shows that there is no significant difference between the two groups.

The contextualized word learning that is represented by created Group 2-4-6 is significantly better than no strategy application. The vocabulary list learning that is represented by created Group 3-5-7 also shows a positive effect, and the effect approaches significance. The contextualized word learning is a little bit better, but not more significant than the vocabulary list learning, while the no strategy application is the worst for this question.

Table 29

*T-test of Question 2 for Analyzing the Contextualized Word Learning*

| Parameter       | Estimate  | Standard Error | t Value | Pr > |t| |
|-----------------|-----------|----------------|---------|------|---|
| 2-4-6 vs. 3-5-7 | 0.02750342| 0.16306713     | 0.17    | 0.8664 |
| 1 vs. 2-4-6     | -0.45061704| 0.22836154    | -1.97   | 0.0508 |
| 1 vs. 3-5-7     | -0.42311362| 0.22949082    | -1.84   | 0.0678 |
As shown in Table 30, for question 3, we can see that the means of created Group 2-4-6 and Group 3-5-7 are higher than the mean of Group 1. The mean of Group 357 is a little higher than the mean of Group 2-4-6. The \( p \) value of Group 1 vs. Group 3-5-7 is 0.0522, which is very close to 0.05, approaching significance. The \( p \) value of Group 1 vs. Group 2-4-6 is relatively low, but the difference between the two groups cannot be considered as significant. The \( p \) value of Group 2-4-6 vs. Group 3-5-7 is much higher than 0.05, which shows that there is no significant difference between the groups.

The contextualized word learning that is represented by created Group 2-4-6 is worse than the vocabulary list learning that is represented by created Group 3-5-7, while the no strategy application is the worst for the third question. The vocabulary list learning is significantly different from no strategy.

Table 30

| Parameter         | Estimate  | Standard Error | t Value | Pr > |t| |
|-------------------|-----------|----------------|---------|-------|---|
| 2-4-6 vs. 3-5-7   | -0.07238595 | 0.11564721     | -0.63   | 0.5326|
| 1 vs. 2-4-6       | -0.24691556 | 0.16195401     | -1.52   | 0.1301|
| 1 vs. 3-5-7       | -0.31930150 | 0.16275490     | -1.96   | 0.0522|

As shown in Table 31, for question 4, we can see that the means of created Group 2-4-6 and Group 3-5-7 are higher than the mean of Group 1. The mean of Group 3-5-7 is a little lower than the mean of Group 2-4-6. All the \( p \) values are much higher than 0.05, which shows that there is no significant difference between the groups. The \( p \) value of Group 2-4-6 vs. Group 3-5-7 is high, which means these two groups are similar.

The contextualized word learning that is represented by created Group 2-4-6 is a little bit better than the vocabulary list learning that is represented by created Group 3-5-7, while the no
strategy application is the worst for this question. However, there is no significant difference between them. The contextualized word learning and vocabulary list learning have similar effects for the fourth question.

Table 31

*T-test of Question 4 for Analyzing the Contextualized Word Learning*

| Parameter       | Estimate | Standard Error | t Value | Pr > |t| |
|-----------------|----------|----------------|---------|------|---|
| 2-4-6 vs. 3-5-7 | 0.01795111 | 0.11621572 | 0.15 | 0.8775 |
| 1 vs. 2-4-6     | -0.14197481 | 0.16275015 | -0.87 | 0.3848 |
| 1 vs. 3-5-7     | -0.12402371 | 0.16355497 | -0.76 | 0.4498 |

As shown in Table 32, for question 5, we can see that the means of created Group 2-4-6 and Group 3-5-7 are higher than the mean of Group 1. The mean of Group 3-5-7 is a little higher than the mean of Group 2-4-6. All the p values are much higher than 0.05, which shows that there is no significant difference between the groups.

The contextualized word learning that is represented by created Group 2-4-6 is a little bit worse than the vocabulary list learning that is represented by created Group 3-5-7, while the no strategy application is the worst for this question. However, there is no significant difference from each other.

Table 32

*T-test of Question 5 for Analyzing the Contextualized Word Learning*

| Parameter       | Estimate | Standard Error | t Value | Pr > |t| |
|-----------------|----------|----------------|---------|------|---|
| 2-4-6 vs. 3-5-7 | -0.08520340 | 0.12748365 | -0.67 | 0.5052 |
| 1 vs. 2-4-6     | -0.03703796 | 0.17852992 | -0.21 | 0.8360 |
| 1 vs. 3-5-7     | -0.12224136 | 0.17941278 | -0.68 | 0.4970 |

As shown in Table 33, for the total score, we can see that the means of created Group 2-4-6 and Group 3-5-7 are higher than the mean of Group 1. The mean of Group 3-5-7 is a little lower than the mean of Group 2-4-6. The p value of Group 1 vs. Group 2-4-6 is 0.0529, which is
very close to 0.05, and shows that the difference between them approaches significance. The \( p \) value of Group 1 vs. Group 3-5-7 is 0.0740, which is a little higher than 0.05, and shows that the difference between Group 1 and Group 3-5-7 approaches significance. The \( p \) value of Group 2-4-6 vs. Group 3-5-7 is much higher than 0.05, which shows that there is no significant difference between the two groups.

The contextualized word learning that is represented by created Group 2-4-6 is significantly better than no strategy application. The vocabulary list learning that is represented by created Group 3-5-7 also shows a positive effect, and the effect approaches significance. The contextualized word learning is a little bit better, but not significantly better than the vocabulary list learning, while the no strategy application is the worst for this question.

Table 33

**T-test of Total Score for Analyzing the Contextualized Word Learning**

| Parameter      | Estimate   | Standard Error | \( t \) Value | \( Pr > |t| \) |
|----------------|------------|----------------|---------------|---------------|
| 2-4-6 vs. 3-5-7| 0.08260082 | 0.41020112     | 0.20          | 0.8408        |
| 1 vs. 2-4-6    | -1.12345648| 0.57445151     | -1.96         | 0.0529        |
| 1 vs. 3-5-7    | -1.04085566| 0.57729226     | -1.80         | 0.0740        |

Table 34 provides a summary of groups of highest and lowest means for each question and total score for contextualized word learning.

Table 34

**Groups of Highest and Lowest Means for Each Question and Total Score for Contextualized Word Learning**

<table>
<thead>
<tr>
<th>Questions</th>
<th>Groups of the Highest Mean</th>
<th>Groups of the lowest Mean</th>
<th>Estimate</th>
<th>( P ) Value of ( t )-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Group 2-4-6</td>
<td>Group 1</td>
<td>0.24691111</td>
<td>0.2458</td>
</tr>
<tr>
<td>2</td>
<td>Group 2-4-6</td>
<td>Group 1</td>
<td>0.45061704</td>
<td>0.0508</td>
</tr>
<tr>
<td>3</td>
<td>Group 3-5-7</td>
<td>Group 1</td>
<td>0.31930150</td>
<td>0.0522</td>
</tr>
<tr>
<td>4</td>
<td>Group 2-4-6</td>
<td>Group 1</td>
<td>0.14197481</td>
<td>0.3848</td>
</tr>
<tr>
<td>5</td>
<td>Group 3-5-7</td>
<td>Group 1</td>
<td>0.12224136</td>
<td>0.4970</td>
</tr>
<tr>
<td>Total</td>
<td>Group 2-4-6</td>
<td>Group 1</td>
<td>1.12345648</td>
<td>0.0529</td>
</tr>
</tbody>
</table>
The order of the means for different questions is relatively consistent, and we can see that Group 2-4-6 is better than the other two groups. Group 2-4-6 is in first place three times for the five questions and has the highest mean for the total score. The differences of means between Group 2-4-6 and Group 1, which is in the last place, are have a difference range from 0.1419 to 0.4506 for the three questions, for which two points are the full points possible for each question. Group 2-4-6 is 1.1235 higher than Group 1, which is in the last place for the total score, and the ten points are the full points. The difference between Group 2-4-6 and Group 1 approaches significance for question 2 and the total score. Therefore, contextualized word learning, which is represented by Group 2-4-6, shows positive effect and the strategy is effective.

Although Group 3-5-7 is a little worse than Group 2-4-6, it also shows a positive effect. Group 3-5-7 is in first place twice for the five questions. For question 2 and the total score, although Group 2-4-6 is in the first place, the mean of Group 3-5-7 is close to the mean of Group 2-4-6, and the difference between Group 1 and Group 3-5-7 approaches significance. Therefore, contextualized word learning, which is represented by Group 2-4-6, is a little better than vocabulary list learning, which is represented by created Group 3-5-7, and both contextualized word learning and vocabulary list learning are better than no strategy, which is represented by Group 1.

Group 1 is in last place six times: for the five questions and the total score. Group 1 is significantly different from the group that is in the first place for question 2, question 3, and the total score. Therefore, any strategy is better than the no strategy application represented by Group 1.
However, the \( p \) values of Group 2-4-6 vs. Group 3-5-7 are much higher than 0.05. Therefore, there is no significant difference between the groups, which means contextualized word learning is not significantly better than vocabulary list learning.

**Other Factors Analysis**

As shown in Table 35, the SAS system also gives us each question’s and total score’s \( p \) values for the model through the GLM procedure, which is a method in SAS. We can see that the \( p \) values of question 3, question 4, question 5, and the total score are lower than 0.05, which means the factors of the model significantly affect the reading comprehension test. However, we saw above that the effect of strategic word learning on reading comprehension is not always significant. Thus, we also need to analyze another factor—how the rater can impact the results of the reading comprehension test in this model.

Table 35

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P ) Values of ( t )-test</td>
<td>0.1763</td>
<td>0.2677</td>
<td>0.0096</td>
<td>0.0002</td>
<td>&lt;.0001</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

As rater A, B, and C evaluated the first forty-nine students and rater D, E, and F evaluated the other seventy-five students, we listed the means for their evaluations separately. For the first part of Table 36, we can see that the means of rater A, B, and C are relatively close for the five questions. For the second part of the table, we can see that the means of rater D, E, and F are relatively different, and the order of the means that the three raters evaluated is not consistent for the five questions. This means that we cannot say that a certain rater is consistently stricter than the others.
Table 36

Means of Different Raters

<table>
<thead>
<tr>
<th>Rater</th>
<th>Question 1</th>
<th>Question 2</th>
<th>Question 3</th>
<th>Question 4</th>
<th>Question 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.8776</td>
<td>1.0204</td>
<td>0.9388</td>
<td>0.6531</td>
<td>1.0204</td>
</tr>
<tr>
<td>B</td>
<td>0.9592</td>
<td>1.0204</td>
<td>0.9796</td>
<td>0.6531</td>
<td>0.8980</td>
</tr>
<tr>
<td>C</td>
<td>0.9184</td>
<td>1.0000</td>
<td>0.8776</td>
<td>0.5510</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rater</th>
<th>Question 1</th>
<th>Question 2</th>
<th>Question 3</th>
<th>Question 4</th>
<th>Question 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>1.1067</td>
<td>1.1600</td>
<td>0.9467</td>
<td>1.4667</td>
<td>1.6933</td>
</tr>
<tr>
<td>E</td>
<td>1.5200</td>
<td>1.2533</td>
<td>1.3467</td>
<td>1.2533</td>
<td>1.5733</td>
</tr>
<tr>
<td>F</td>
<td>1.2533</td>
<td>1.2133</td>
<td>1.0267</td>
<td>0.7467</td>
<td>1.4800</td>
</tr>
</tbody>
</table>

As shown in Table 37, rater A, B, and C have relatively high correlation for question 1. Rater D, E, and F also have high correlation for this question. However, the correlations among rater D, E, and F are a little lower than the correlations among the first three raters.

Table 37

Pearson Correlation Coefficients of Raters for Question 1

<table>
<thead>
<tr>
<th></th>
<th>Rater A</th>
<th>Rater B</th>
<th></th>
<th>Rater D</th>
<th>Rater E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater B</td>
<td>0.7713</td>
<td>0.8370</td>
<td></td>
<td>Rater E</td>
<td>0.8120</td>
</tr>
<tr>
<td>Rater C</td>
<td>0.9235</td>
<td>0.8968</td>
<td>Rater F</td>
<td>0.7153</td>
<td>0.7482</td>
</tr>
</tbody>
</table>

As shown in Table 38, rater A, B, and C have relatively high correlation for question 2. Rater D, E, and F also have high correlation for this question. However, the correlations among rater D, E, and F are a little lower than the correlations among the first three raters.

Table 38

Pearson Correlation Coefficients of Raters for Question 2

<table>
<thead>
<tr>
<th></th>
<th>Rater A</th>
<th>Rater B</th>
<th></th>
<th>Rater D</th>
<th>Rater E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater B</td>
<td>0.9050</td>
<td>0.8714</td>
<td></td>
<td>Rater E</td>
<td>0.8202</td>
</tr>
<tr>
<td>Rater C</td>
<td>0.8968</td>
<td>0.8714</td>
<td>Rater F</td>
<td>0.8004</td>
<td>0.8189</td>
</tr>
</tbody>
</table>
As shown in Table 39, rater A, B, and C have relatively high correlation for question 3. Rater D, E, and F have low correlation for this question. Rater F is especially different from raters D and E.

Table 39

*Pearson Correlation Coefficients of Raters for Question 3*

<table>
<thead>
<tr>
<th></th>
<th>Rater A</th>
<th>Rater B</th>
<th>Rater D</th>
<th>Rater E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater B</td>
<td>0.8727</td>
<td></td>
<td></td>
<td>0.6426</td>
</tr>
<tr>
<td>Rater C</td>
<td>0.9039</td>
<td>0.8621</td>
<td></td>
<td>0.5495</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rater F</td>
<td>0.5445</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As shown in Table 40, rater A, B, and C have relatively high correlation for question 4. Rater D, E, and F have low correlation for this question. Rater F is especially different from rater D.

Table 40

*Pearson Correlation Coefficients of Raters for Question 4*

<table>
<thead>
<tr>
<th></th>
<th>Rater A</th>
<th>Rater B</th>
<th>Rater D</th>
<th>Rater E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater B</td>
<td>0.7549</td>
<td></td>
<td></td>
<td>0.5917</td>
</tr>
<tr>
<td>Rater C</td>
<td>0.9112</td>
<td>0.7884</td>
<td></td>
<td>0.2221</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rater F</td>
<td>0.4669</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As shown in Table 41, rater A, B, and C have relatively high correlation for question 5. Rater D, E, and F also have relatively high correlation for this question, but the correlations among rater D, E, and F are lower than the correlations among the first three raters.

Table 41

*Pearson Correlation Coefficients of Raters for Question 5*

<table>
<thead>
<tr>
<th></th>
<th>Rater A</th>
<th>Rater B</th>
<th>Rater D</th>
<th>Rater E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater B</td>
<td>0.8266</td>
<td></td>
<td></td>
<td>0.7835</td>
</tr>
<tr>
<td>Rater C</td>
<td>0.9496</td>
<td>0.8085</td>
<td></td>
<td>0.7220</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rater F</td>
<td>0.6988</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In short, the first three raters, A, B, and C, have a high correlation for the five questions, which means they have similar evaluations for the students’ answers. The second three raters
have a relatively high correlation for question 1, 2, and 5. However, rater F has a very different evaluation from rater D and E for question 3 and question 4. Rater D and E also have relatively low correlation for question 4, which means their evaluations are not very similar for the question. Thus, rater reliability is a factor that influenced the test’s results, which is one reason that the p values for the model are significant.

Summary

The statistical analyses of this experiment indicate that the combination of contextualized word learning, pre-learning, and normal learning is the most effective among the six combined word learning strategies but is not statistically significant. Considering pre-learning, the effect of pre-learning combined with normal learning is statistically significant and therefore the strategy is effective. The effect of pre-learning alone is very small, and pre-learning is almost the same as the no strategy application, both not significant. The contextualized word learning shows a positive effect, but it is not significantly better than vocabulary list learning. The second three raters have a low correlation for question 3 and question 4, which also impacts the results of the test.
Chapter Five: Discussion and Conclusions

Overview of the Study

The aim of this study is to examine whether pre-learning vocabulary instruction and contextualized word learning affect Chinese reading comprehension for intermediate Chinese foreign language learners. There are four vocabulary learning strategies in the study: contextualized word learning, vocabulary list learning, pre-learning, and normal-learning. In the experiment we used six treatments that combined the four different vocabulary learning strategies in order to compare the effects of pre-learning and normal-learning and the effects of contextualized word learning and vocabulary list learning on reading comprehension.

To implement the experiment, the subjects of different groups read the same reading passage and then answered the same five questions without the reading passage or any other references. The control group (Group 1) did not receive any vocabulary to learn during the process of reading. The treatment groups learned vocabulary in six ways: Group 2, which represents combining contextualized word learning and pre-learning, read a vocabulary list with example sentences before receiving the reading passage. Group 3, which represents combining vocabulary list learning and pre-learning, read a vocabulary list before receiving the reading passage. Group 4, which represents combining contextualized word learning, pre-learning, and normal-learning, read a vocabulary list with example sentences before receiving the reading passage, and participants were allowed to refer to the list and sentences while reading the passage. Group 5, which represents combining vocabulary list learning, pre-learning, and normal-learning, read a vocabulary list before receiving the reading passage and they were allowed to refer to the list while reading the passage. Group 6, which represents combining
contextualized word learning and normal-learning, received a vocabulary list with example sentences and the reading passage at the same. Group 7, which represents vocabulary list learning and normal-learning, received a vocabulary list and the reading passage at the same time.

The study discovered that the combination of contextualized word learning, pre-learning, and normal-learning showed significant positive effects on reading comprehension. Pre-learning itself without normal learning did not show significant positive effects on reading comprehension, while normal learning is a little better than no-strategy learning for understanding. Contextualized word learning shows positive effects, but it is not significantly different from vocabulary list learning and no-strategy learning. The results also show that the evaluations of the six raters did not always have a high correlation, which is a factor that we need to consider while analyzing the results of the reading comprehension tests.

**Discussion of the Results**

**Discussion of Basic Data Analysis.** From the basic information, we can see that Group 4, which represents a combination of contextualized word learning, pre-learning, and normal learning, scored the highest although is not statistically significant. This is a reasonable result. The first reason is that the students of Group 4 had the most time to learn (i.e. both the pre-learning period when they studied the vocabulary before reading the passage and the normal learning period when they could continue to use the vocabulary explanation while reading the passage). When they read the vocabulary list with example sentences before reading the passage, they had background information of the new words. When they went back to the vocabulary list after initially reading the passage, they continued to learn the words in context. At the same time, they used the knowledge of the new words to understand the passage. The second reason is that the students of Group 4 received the most information. In their vocabulary learning sheet, they
had the same English translation as the other groups’ normal vocabulary lists, and they also had example sentences that can help them understand the words more deeply, which helps them to understand the words in a context and understand the passage with the deeper understanding of the words.

The means of Group 1 (the control group), Group 2 (combination of contextualized word learning and pre-learning vocabulary), and Group 3, (combination of vocabulary list learning and pre-learning) were close and they had the three lowest scores and were not statistically significant. The students of Group 1 did not receive any treatment. They did not have resources to learn any vocabulary before or during the process of reading. They understood the passage based solely on the knowledge that they already had. It is normal and expected that Group 1 had the lowest score for most questions.

However, despite the treatments that Group 2 and Group 3 received, they scored close to Group 1 and sometimes even worse than Group 1, which is unexpected considering that both Group 2 and Group 3 had pre-learning. The students of Group 2 had contextualized word learning in which they had example sentences that could help them understand the words deeply. However, Group 2 scored a little worse than Group 3, which had vocabulary list learning without example sentences. The following may be reasons that explain the results. Group 2 and Group 3 only had five minutes to learn these new words, but the students of Group 2 had more information to read than Group 3 as their vocabulary learning sheet included example sentences. The students of Group 2 needed to learn more than the students of Group 3 with limited time, therefore the students of Group 2 might have understood the information in the vocabulary learning sheet, but they did not have enough time to remember it all, while the students of Group 3 could have used their limited time to look at the vocabulary list several times and simply
memorize the meanings, which could help them understand the passage. With limited time, it is more difficult for learners to understand and remember lots of information. Therefore, in this experiment, Group 2 who used pre-learning and contextualized word learning is a little worse than Group 3 who used pre-learning and vocabulary list learning. Thus, with limited time, contextualized word learning may not show its potential power, but with more time, contextualized word learning may be more effective. The combinations of strategies that Group 2 and Group 3 used did not show any strong positive effect, and one reason may be that the students were not used to using these strategies.

Except for Group 4 remaining relatively consistent in first place, the orders of the groups are different for the five questions and the total score. One reason may be that the number of participants was relatively limited. It is possible that the orders of different questions may be relatively consistent when we have more students to participate in the experiment. The second reason is that the students’ Chinese language abilities are different, although all of them were in Chinese 301 when they took the test. It is possible that few students already knew most of the words in the passage, and they could completely understand the passage without learning any words, while few other students could not understand the passage even though they had the vocabulary translation. Also, it is possible that some students’ reading speed is slower, and reading example sentences of the new vocabulary took up some of the time they needed to understand the passage. Thus, the results may reflect not only the treatments’ effects but also a part of the students’ initial Chinese language abilities.

Discussion of Research Question One. This question asks: *Does pre-learning vocabulary instruction affect CFL (Chinese Foreign Language) learners’ reading comprehension?*
As the Research Question One Analysis section in Chapter Four mentioned, Group 4-5 (a combination of Group 4 and Group 5 to represent pre-learning and normal-learning) is relatively consistent in placing at the top and is statistically significant. Group 2-3 (a combination of Group 2 and Group 3 to represent pre-learning) and Group 1 (the control group) were very close and are the last groups, although Group 2-3 was in the first place for question 2. The results imply that pre-learning with normal learning can have a significant positive effect on reading comprehension, while sometimes pre-learning without normal learning is actually worse than just normal learning which is represented by Group 6-7 (a combination of Group 6 and Group 7). Pre-learning itself did not show a positive effect on learners’ reading comprehension.

Pre-learning with normal learning positively affects reading comprehension (statistically significant), which is reasonable, because the students had more time to learn new words. They learned the vocabulary before reading the passage and they had access to the vocabulary explanation while reading the passage.

The possible reasons that pre-learning without normal learning is worse than just normal learning are as follows: Chinese language pre-learning is improved with better memorization ability, which may require long-term training to obtain. The experiment offers only one chance for the students who were in Group 2 and Group 3, so it may be relatively difficult for the students who were not used to using the pre-learning strategy to memorize the meaning of the new words in the limited time, despite understanding the words during the five minutes of pre-learning. The reason that pre-learning itself did not show a positive effect on reading comprehension may also be that pre-learning is aided by more training.

Second, when the students of Group 6-7 (representing normal learning) read the passage and the vocabulary learning sheet at the same time, they could completely understand the words
through the context in the passage, which is the same as the function of contextualized word learning. The understanding of the words through the context in the passage directly helps the students understand the meaning of the passage, which indirectly shows that contextualized word learning has positive effects on reading comprehension.

Third, pre-learning may not be as helpful for learning Chinese vocabulary, especially without any training, which is unexpected. However, as I mentioned in Chapter Two, because of the characteristics of Chinese language and Chinese characters, understanding a Chinese character or word depends more on context than other languages. Thus, simply memorizing translations of new Chinese vocabulary before reading a passage, which was represented by Group 3, may not be very useful to understand the context of the passage. Group 2 had sample sentences; however, they had limited time and relatively simple examples in this experiment, and it might be difficult for Group 2 to memorize lots of information and then apply that information in reading, especially without training in class.

For research question one, we found that the effect of pre-learning is very small, and the strategy is not as effective for learning Chinese vocabulary for reading comprehension, which may be due to the characteristics of Chinese language and a lack of training in using the strategy.

Discussion of Research Question Two. This question asks: Does teaching contextualized word learning strategy affect CFL learners’ reading comprehension?

As the Research Question Two Analysis section in Chapter Four mentioned, Group 2-4-6 (a combination of Group 2, Group 4, and Group 6 to represent contextualized word learning) is relatively consistent in placing at the top. Group 1 (the control group) is consistently in the last place. For question 2 and the total score, Group 2-4-6 is significantly different from Group 1, which is the control group. The results imply that the contextualized word learning strategy has a
positive effect on reading comprehension, which is expected. However, sometimes the effect of contextualized word learning is not statistically significant in this experiment, which was an unexpected result.

Contextualized word learning is the process by which learners study new vocabulary with example sentences. The examples sentences that were made with the new words help learners understand the words more completely, so that the learners can understand the meaning of a word in a specific context, which is good for the learners’ understanding of a passage. Also, through the examples sentences, learners can know how to use the new words when they create their own sentences or passage, which also indirectly help learners improve their reading comprehension ability. Thus, that contextualized word learning has a positive effect on reading comprehension is reasonable.

This also implies that contextualized word learning corresponds with the characteristics of Chinese language. As we mentioned in Chapter Two, understanding through context is very important for learning Chinese characters and improving Chinese reading comprehension. For some sentences, foreigners may not think they are clearly written, while native Chinese speakers can completely understand them. The reasons may be that, first, the meanings of some words change in different contexts, and second, the part of speech of some words also changes in different contexts. For example, “守护天使 shǒuhù tiānshì” can be understood as “to protect or protected a person or people who is or are like an angel or angels,” and it also can be understood as, “an angel or angels who is or are guardian or guardians”. The meaning depends on the context.
As an example from this study, the answer to the second question is *peaceful* in English; however, in other situations, the translation may be *in peace*. Although we gave the English translation as an adjective, we still believe that with examples sentences learners can understand more, which may be one reason that for question 2 contextualized word learning shows a significantly positive effect.

However, in the experiment the positive effect of contextualized word learning was not significant for the other questions. The analysis follows. First, the number of participants is relatively limited. One hundred twenty-four students participated in the experiment. Group 3 and Group 5 each had seventeen students, and the other groups each had eighteen students. These figures met the requirements for statistical analysis. However, the students’ initial reading and learning abilities and Chinese experiences are variable; therefore, the number of the subjects is still limited in order to balance the differences of background. If we could have had more participants, maybe we would see more significant positive results in the other questions.

Second, as mentioned above, students’ initial reading ability and Chinese knowledge may have impacted the results of the experiment. Some students who had example sentences in their vocabulary learning sheet may have a weaker reading ability. So, they might also have difficulties understanding the passage although they learned the new words with example sentences, because they also needed to learn the example sentences, which likely consumed more of their time and energy. Other students who were in Group 1 (the control group) may have understood the passage without any learning because of their previous knowledge. Thus, the effect of contextualized word learning may be influenced by the factor of students’ initial Chinese knowledge. This problem can be solved by having more participants and by screening for prior knowledge and skills.
Third, for some students whose Chinese abilities are weaker, it may take more time for them to learn how to study new words with a contextualized word learning strategy. As mentioned above, reading the example sentences consumes their time and energy. Therefore, a long-term training for learning the strategy may be good for the learner, especially for relatively low-level students.

From the results of research question two, we also can see that Group 3-5-7 is in first place for question 3 and question 5, and for question 3, the difference with Group 1 approaches statistical significance. For the other questions and the total score, Group 3-5-7 is better than Group 1. Thus, vocabulary list learning has a positive effect on reading comprehension, although it is a little worse than contextualized word learning and the positive effect is not significant in this experiment.

The answer to question 3 is a verb, and its meaning does not change much in different situations. Thus, simply memorizing the translation, which was represented by Group 3-5-7, may be good for learning some Chinese verbs, while contextualized word learning gave more information that may be difficult for learners to understand and remember with limited time. This may be one reason that Group 3-5-7 is in the first place for question 3 and the difference from Group 1 approaches significance.

The answer to question 5 is a verb as well. Group 3-5-7 is also better than Group 2-4-6 and Group 1 for this question. However, the difference is not significant. One reason may be that this verb changed part of speech in the context, for which contextual understanding may be better than knowing only the English translation.

Although contextualized word learning is a little better than vocabulary list learning, the difference is not significant, and for question 2, question 4, and the total score, the $p$ values of the
t-tests for Group 2-4-6’s vs. Group 3-5-7, which represented contextualized word learning and vocabulary list learning respectively, are very high, which means the effects of these two strategies are close for these two questions and the total score. It implies that both strategies are better than not applying any strategy. For certain words, such as some verbs, contextualized word learning may not show its power, and for some other kinds of words, such as adjectives, it may be more effective.

For research question two, we found that contextualized word learning has a positive effect on Chinese reading comprehension. However, it may not fully show its effect for certain words.

**Discussion of Other Factors.** As the Other Factors Analysis section in Chapter Four mentioned, the means of the raters’ evaluations are different. The first three raters and the second three raters evaluated different students’ answers. We do not analyze the difference between these two groups, because this study focuses on students’ performances, and students are different for the two groups’ raters, which can be one reason that the evaluations of the two groups are different. The first three raters’ evaluations are relatively similar. One reason may be that rater B and rater C knew rater A’s evaluation when they evaluated the answers, thus rater A’s thoughts might have influenced rater B and rater C.

The second three raters did not know the others’ evaluations, so they did not influence each other. However, their evaluations are relatively different, especially question 4 where the correlation coefficient of rater D and rater F are very low. One reason is that they may have understood the students’ English answers differently. For some English answers, one rater might think it was close to the right meaning that shows the student understood the word, while another rater might think it was not the right meaning that shows the student did not understand the word
but might understand the context. The second reason is that the raters may have had slightly different understandings of the target words, thus for some students’ translation of the target words, one rater might think it was right, while the other one might think it was only close to the right meaning.

The problem is the understanding of the first language and the second language not only for the raters but also for the students, which relates to two languages’ translation. This study does not focus on this problem, which better training would have solved, thus we tried to use the three raters to balance the different situations.

For question 4, the direct English translation of the answer is *frequency*. However, in general, English native speakers do not use the word in the context that the Chinese sentence in the passage tries to express. In order to answer the question, the students who did the test needed to explain the sentence in English, which the target word is in. Thus, the raters have more room to judge students’ answers, which can be one reason that rater D and F have very low correlation.

Also, students’ understanding of the five questions may also influence the results of the experiment. As the raters discussed the five questions with the researcher, they understood the researcher’s intention. However, some students might have questions about the five questions of the test, as they are not supposed to know what the study wanted to test. So, for some students’ answers, a rater might have thought it did not relate to the question, while the other rater might have been able to understand the students’ thought process after considering the question and gave one point for the answer. This problem cannot be solved in this study, as we cannot use different questions for a single experiment for the different semesters. The researcher should carefully reconsider the questions for future research.
Implications of the Results

From the results of this study there are several implications in terms of Chinese language acquisition and reading comprehension. Group 4 who had pre-learning, normal learning, and contextualized word learning showed a significantly positive difference from other groups. We can see that learning Chinese takes time and requires background knowledge and context information, which reflects the characteristics of the Chinese language that we analyzed in Chapter Two.

Contextualized word learning has positive effects on reading comprehension, so we need to encourage students to learn Chinese vocabulary in context, which can show the meaning of the word in a specific situation. This can help learners understand new words more deeply and know how to use the new words. Only memorizing the English translation of new Chinese words is not as helpful.

However, pre-learning alone did not show a positive effect. Also, although contextualized word learning shows the positive effect on reading comprehension, it is not always significant. One possible reason for this is that the treatments that were used in the experiments should have been used for a longer time period. These results might imply that learning Chinese language and improving Chinese reading comprehension is a long-term process. Pre-learning and contextualized word learning may be good strategies, but training learners to learn how to use the strategies requires time. Using the strategies effectively is an asset in learning the Chinese language. Learners may obtain the ability gradually through practice.
Limitations of this Study

One limitation of this study was that the number of subjects in each group was relatively small. This prevented the experiment from fully showing the effects of pre-learning and contextualized word learning strategies on Chinese reading comprehension.

Another limitation of the current study was that the questions in the test could have been more precise. From the students’ answers, we can see that a small number of students either did not understand or misunderstood the questions. For example, a few students thought “why” in question #1 meant “where”. It would be better in the future to underline the key question words. Also, the English words in the questions should be considered more carefully and modified so that the questions could clearly point out which part of the passage is in question. For example, question #5 asked for the sound that the author heard after leaving, but a few students answered the second sound that the author heard, which led to disagreement among the raters. One rater thought that the second sound was also heard by the author after he left, so the answer could be partly right, while another rater thought the answer was totally wrong, as they might not understand the first sound.

Insufficient training of the raters is a limitation of this study as well. We should let all the raters discuss together before starting their evaluations instead of letting rater B and rater C read the evaluations of rater A, to avoid having rater A influence the other two raters. Also, by discussing what they thought about possible controversial answers before starting to evaluate, they would come to a consensus about how to deal with controversial answers.

Recommendations for Future Research

Since the strategies of pre-learning and contextualized word learning can be gradually practiced so that learners can obtain the abilities to use the strategies in reading comprehension, a
future study may design a long-term experiment, such as one semester long. The experiment could include a pre-test and a post-test. During the interim period, one treatment group can have regular training in pre-learning strategizing, and another treatment group can have regular training in contextualized word learning strategizing. Through comparing the results of the pre-test and the post-test, we could more clearly see the effect of the two strategies on reading comprehension, because the results of the pre-test should be the same as the experiment in the current study, and the difference between the pre-test and the post-test could indicate the influence of students’ initial reading abilities. The design of a future experiment may also show that through long-term training learners may obtain skills of using pre-learning and contextualized word learning that are good for their Chinese reading comprehension and Chinese language learning.

Another recommendation for future research would be to analyze the way of learning different types of words. For example, in Chinese, learning verbs may be relatively easy for American students, as they do not need to be conjugated. However, learning Chinese adjectives may be relatively difficult, because there are so many adjectives in Chinese, and in general they have different meanings in different situations or contexts. So, contextualized word learning may be more appropriate for adjectives, and vocabulary list learning may be sufficient for learning Chinese verbs and some nouns.

A future study may also consider Chinese vocabulary learning for reading comprehension at different levels. Contextualized word learning may be better for advanced-level students, because at this level the meaning of one word may depend mainly on the context, and reading comprehension can identify students’ comprehensive abilities, which include students’ background knowledge. Thus, training learners in the ability to get more detailed information
and understand different meanings of one word in different contexts is a good way to improve their Chinese reading. Vocabulary list learning may be good for beginning learners, because at this level learners focus on learning simple nouns and verbs that do not vary much in different situations. Also, at this level learners have limited vocabulary, and it is difficult for them to learn new words through the words that they learn.

This research studied contextualized word learning, which showed positive effects on reading comprehension, and pre-learning, which did not show positive effect in this experiment. This study also analyzed some reasons to explain the results. With the analysis of this experiment’s results, we also gave some suggestions for future research.
References


The Defense Language Institute Foreign Language Center’s office of Strategic Communications.


天津市市内六区教研室联合编写. (2009). 小学学习测评，语文，三年级下人教版。天津人民出版社，天津市西康路 35 号，邮编 300051。


Appendix A: Consent Form I

Identifying Successful Reading Strategies Among Intermediate / Advanced learners of Mandarin Chinese

Introduction

This research study is being conducted by Hongyi Jia at Brigham Young University to determine what methods of learning Chinese vocabulary are good for students to improve their reading comprehension. You were invited to participate because you are currently taking a Chinese 301 class.

Procedures

You do usual reading comprehension exercises in your normal classes. Half of the students learn vocabulary before they read Chinese passages, and the other half learn vocabulary when they read Chinese passages or after they read Chinese passages. Your scores from these exercises will be used for research purposes.

Risks/Discomforts

There are minimal risks in this study.

Benefits

The results of the study will be direct recommendations for BYU’s Chinese study program. It could conceivably benefit Chinese learners, BYU, and Chinese language acquisition researchers by providing data regarding factors facilitating Chinese reading comprehension.

Confidentiality

All information provided will remain confidential and will only be reported as group data with no identifying information. All data will be kept in a storage cabinet and only those directly involved with the research will have access to them. After the research is completed, the data will be destroyed.

Compensation

Participants will receive 10 extra credit points in Chinese 301 for consenting that the results of their usual reading comprehension exercises in their normal class can be analyzed in the study. For those who do not wish to participate in the research, 10 extra credit points can be earned by doing other Chinese activities held by BYU.

Participation
Participation in this research study is voluntary. You have the right to refuse to participate the study without jeopardy to your class status, grade or standing with the university.

Questions about the Research
If you have questions regarding this study, you may contact Hongyi Jia, at (646) 715-4668, jhygeneva@yahoo.com

Questions about your Rights as Research Participants
If you have questions regarding your rights as a research participant, you may contact the BYU IRB Administrator at (801) 422-1461, A-285 ASB, Brigham Young University, Provo, UT 84602, irb@byu.edu.

I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.
Signature: ____________________________________________ Date: ____________
Appendix B: Consent Form II

Identifying Successful Reading Strategies Among Intermediate / Advanced learners of Mandarin Chinese

Introduction

This research study is being conducted by Hongyi Jia at Brigham Young University to determine what methods of learning Chinese vocabulary are good for students to improve their reading comprehension. You were invited to participate because you are currently taking a Chinese 301 class.

Procedures

You do extra reading comprehension exercises after class. Half of the students learn vocabulary before they read Chinese passages, and the other half learn vocabulary when they read Chinese passages or after they read Chinese passages. Your scores from these exercises will be used for research purposes.

Risks/Discomforts

There are minimal risks in this study.

Benefits

The results of the study will be direct recommendations for BYU’s Chinese study program. It could conceivably benefit Chinese learners, BYU, and Chinese language acquisition researchers by providing data regarding factors facilitating Chinese reading comprehension.

Confidentiality

All information provided will remain confidential and will only be reported as group data with no identifying information. All data will be kept in a storage cabinet and only those directly involved with the research will have access to them. After the research is completed, the data will be destroyed.

Compensation

In appreciation for students’ consent that their extra reading comprehension exercises can be analyzed in my thesis, students will get extra help for their Chinese.

Participation
Participation in this research study is voluntary. You have the right to refuse to participate the study without jeopardy to your class status, grade or standing with the university.

Questions about the Research
If you have questions regarding this study, you may contact Hongyi Jia, at (646) 715-4668, jhygeneva@yahoo.com

Questions about your Rights as Research Participants
If you have questions regarding your rights as a research participant, you may contact the BYU IRB Administrator at (801) 422-1461, A-285 ASB, Brigham Young University, Provo, UT 84602, irb@byu.edu.

I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.
Signature:_________________________________________ Date:__________
马大姐开车

马大姐开车

马大姐走到哪儿就会把笑声带到哪儿。别的不说，就说马大姐学开车这件事儿吧，您可能想不到，她十年前就拿到了汽车驾驶证，也算是“老司机”了，可是她只开过两次汽车。第一次是拿到驾驶证的当天晚上。那天，下着小雨，她把车开到天安门广场。她一看见红灯就着急了，找不着停车线了。她的车正在走走停停的时候，警察发现了她，大声地叫喊：“靠边停车！”马大姐马上把车开到路边停住，然后下车跑到警察身边，很恭敬地对警察说：“对不起！您要罚多少？”警察问她：“红灯都亮了，你还往前开！你要去哪儿？”马大姐觉得自己是第一天开车，犯点儿小错，警察应该原谅她。可是警察看都没看她，就说：“十块钱！”马大姐把驾驶证和钱递过去，警察拿起驾驶证一看：“是您啊，马大姐。”然后给她一张罚款单，说：“您去银行交钱，下次请注意！”

第二次开车出去，刚到十字路口，车就熄火了，怎么也发动不起来。警察走过来问：“您在这儿干嘛呢？”“没干什么，我看见警察有点儿紧张！”“又没犯错误！您紧张什么？用不用帮您推一推？”马大姐说：“不麻烦您了！”

十多年来，马大姐就开了这么两次车，而且两次都很不顺利。后来她再也没有开过车。
Appendix D  mǎdàjiě kāichē 马大姐开车 Questions

Questions

1. Why did she get the first ticket?

2. What tone did she use with the first policeman?

3. How did she think the first policeman should have treated her?

4. The second time she was driving, what happened to her car?

5. Why did the second policeman say she shouldn’t be nervous?
Appendix E: 马大姐开车 Vocabulary List with Example

广场: public square
例句: 去纽约(New York)的游客(tourists)都要去时代广场(Times Square)。

红灯: traffic light
例句: 超红灯(run a red light)的时候不要让警察看到。

停车线: the line at a traffic light you must stop behind.
例句: 红灯的时候车一定要停在停车线后。

恭敬: respectfully
学生要非常恭敬地对老师说话。

罚: punishment, fine
乱丢垃圾(litter)就会被警察罚钱。

原谅: forgive
别人对你不好，你还要原谅他们。

递: handover
把作业递给老师。

罚款单: ticket
要把罚款单和钱一起寄(mail)到法庭(court)。

熄火: extinguish; (of an engine) die
我的车在十字路口(intersection)中间熄火了。

犯错误: make a mistake
在301的课上，表演的时候犯错误不一定要扣(kòu)分。
Appendix F: mǎdàjiě kāichē 马大姐开车 Vocabulary List

广场: public square

红灯： traffic light

停车线： the line at a traffic light you must stop behind.

恭敬： respectfully

罚： punishment, fine

原谅: forgive

递: handover

罚款单: ticket

熄火： extinguish; (of an engine) die

犯错误: make a mistake
Appendix G: zhōngguó kējǔ zhìdù 中国科举制度

中国传统文化有近五千年的发展史，其中以孔子为代表的儒家思想起了重要的作用。由于儒家思想的影响，自古中国很重视学习。隋朝 (Sui Dynasty 581-618) 的时候就发明了一个叫“科举”的考试制度。只有在科举中考了及格的人才能当官。这是世界上第一个公务员考试，而此发明让中国更接近精英领导体制。当时的中国人如果能考过科举当官，他们的生活条件和社会地位一瞬间就会有巨大的改善，所以人人都以此为目标。为了达到此目标，父母会辛苦地攒钱给儿子上学，还会让孩子拼命读书。这种科举体制给中国人留下的思想就是，教育不仅能提高文化修养，也是有实际用处的。

中国现在的社会情况对中国人的教育观念也有很大的影响。现在已经没有科举了，但有高考来做替代。只有考过高考的学生才能进入大学，而只有大学毕业的人才能找到好工作。没有上过大学的人大部分不得不去工厂干很辛苦的活。现在的中国父母都清楚地知道教育是致富之路，是解脱贫困的最可靠的方法。中国父母之所以会给孩子很多压力让他们认真地读书，花钱送他们到补习班或给他们请家教，不只是因为教育本身是一件好事，而是因为他们确信只有这样他们的孩子将来才能过一个好的生活。

中国人对教育的看法受了近五千年的文化和历史的影响，而中国现在的实际情况又是在增强这个观念。教育显然是一件极为重要的事，也是对所有人的生活都会有利的，但中国家长给孩子的这种压力，让他们上完九个小时的课就上补习班，上完补习班就回家熬夜写功课或做习题，不仅对孩子的成长发育不利，也是对他们的身体健康很不好的。我希望中国将来能保持这种重视教育的思想，而且能找到一个比较健全的教育方式。
中國傳統文化有近五千年的發展史，其中以孔子為代表的儒家思想起了重要的作用。由於儒家思想的影響，自古中國很重視學習。隋朝（Sui Dynasty 581-618）的時候就發明了一個叫“科舉”的考試制度。只有在科舉中考了及格的人才能當官。這是世界上第一個公務員考試，而此發明讓中國更接近精英領導體制。當時的中國人如果能考過科舉當官，他們的生活條件和社會地位一瞬間就會有巨大的改善，所以人人都以此為目標。為了達到此目標，父母會辛苦地攢錢給兒子上學，還會讓孩子拼命讀書。這種科舉體制給中國人留下的思想就是，教育不僅能提高文化修養，也是有實際用處的。

中國現在的社會情況對中國人的教育觀念也有很大的影響。現在已經沒有科舉了，但有高考來做替代。只有考過高考的學生才能進入大學，而只有大學畢業的人才能找到好工作。沒有上過大學的人大部分不得不去工廠幹很辛苦的活。現在的中國父母都清楚地知道教育是致富之路，是解脫貧困的最可靠的方法。中國父母之所以會給孩子很多壓力讓他們認真地讀書，花錢送他們到補習班或給他們請家教，不只是因為教育本身是一件好事，而是因為他們確信只有這樣他們的孩子將來才能過一個好的生活。

中國人對教育的看法受了近五千年的文化和歷史的影響，而中國現在的實際情況又是在加強這個觀念。教育顯然是一件極為重要的事，也是對所有人的生活都會有利的，但中國家長給孩子的這種壓力，讓他們上完九個小時的課就上補習班，上完補習班就回家熬夜寫功課或做習題，不僅對孩子的成長發育不利，也是對他們的身體健康很不好的。我希望中國將來能保持這種珍視教育的思想，而且能找到一個比較健全的教育方式。
Appendix H: zhōngguó kējǔ zhìdù 中国科举制度 Questions

1. What benefits did the Imperial Examination (科举) system have for China?

2. How did ancient Chinese parents approach their children's education?

3. Besides improving one’s material situation, why else did the ancient Chinese value education?

4. What problem do modern Chinese parents feel education is a solution to?

5. What does the author say modern Chinese parents make their children do which is bad for their health?
Appendix I: zhōngguó kējǔ zhìdù 中国科举制度 Vocabulary List with Example

儒家 – The Confucian school of thought
儒家思想是孔子(Confucius) 开创的。

科举 – The Imperial Examination system, used in China from the Sui Dynasty till the Qing Dynasty (~581-1911)
科举考试的内容是四书五经(The four books and five classics of Confucian thought)。

精英领导体制 – Meritocracy
精英领导体制是用最优秀(outstanding, excellent) 的人做领导的体制。

一瞬间 – In an instant, in the twinkling of an eye
你不能一瞬间便成百万富翁(millionaire)，需要长时间的努力(effort, hard work)。

攒钱 – Save money
在美国，买房子之前需要攒几万块钱。

拼命 – Exert one’s utmost, desperately
你拼命地工作就有机会升级(be promoted)。

修养 – Self-cultivation; accomplishment, mastery, taste (in art etc.)
很多人认为喜欢看Jerry Springer的节目(show, program)就说明文化修养不高。

替代 – Substitute for, replace, supersede
现在科学家在研究(research)新的替代能源(alternative energy sources)。

解脱 – Escape, release, free oneself, extricate oneself
期末考试(finals)考完了以后有一种解脱的感觉。

贫困 – Poverty
如果一个人在贫困的家庭长大(grow up)，他可能会比较节省(thrifty)。

熬夜 – Burn the midnight oil, pull an all-nighter
如果我今天晚上不熬夜写功课(homework)就写不完。

珍视 – To value, to treasure, to prize
后期圣徒教会(LDS Church)非常珍视家庭。
儒家 – The Confucian school of thought
儒家思想是孔子(Confucius)開創的。

科挙 – The Imperial Examination system, used in China from the Sui Dynasty till the Qing Dynasty (~581-1911)
科挙考試的內容是四書五經(The four books and five classics of Confucian thought)。

精英領導體制 – Meritocracy
精英領導體制是用最優秀(outstanding, excellent)的人做領導的體制。

一瞬間 – In an instant, in the twinkling of an eye
你不能一瞬間便成百萬富翁(millionaire)，需要長時間的努力(effort, hard work)。

攢錢 – Save money
在美國，買房子之前需要攢幾萬塊錢。

拼命 – Exert one’s utmost, desperately
你拼命地工作就有機會升級(be promoted)。

修養 – Self-cultivation; accomplishment, mastery, taste (in art etc.)
很多人認為喜歡看Jerry Springer的節目(show, program)就說明文化修養不高。

替代 – Substitute for, replace, supersed
現在科學家在研究(research)新的替代能源(alternative energy sources)。

解脫 – Escape, release, free oneself, extricate oneself
期末考試(finals)考完了以後有一種解脫的感覺。

貧困 – Poverty
如果一個人在貧困的家庭長大(grow up)，他可能會比較節省(thrifty)。

熬夜 – Burn the midnight oil, pull an all-nighter
如果我今天晚上不熬夜寫功課(homework)就寫不完。

珍視 – To value, to treasure, to prize
後期聖徒教會(LDS Church)非常珍視家庭。
Appendix J: zhōngguó kējǔ zhì 中国科举制度 Vocabulary List

儒家 – The Confucian school of thought

科举 – The Imperial Examination system, used in China from the Sui Dynasty till the Qing Dynasty (~581-1911)

精英领导体制 – Meritocracy

一瞬间 – In an instant, in the twinkling of an eye

攒钱 – Save money

拼命 – Exert one’s utmost, desperately

修养 – Self-cultivation; accomplishment, mastery, taste (in art etc.)

替代 – Substitute for, replace, supersede

解脱 – Escape

贫困 – Poverty

熬夜 – Burn the midnight oil, pull an all-nighter

珍视 – To value, to treasure, to prize
儒家 – The Confucian school of thought

科舉 – The Imperial Examination system, used in China from the Sui Dynasty till the Qing Dynasty (~581-1911)

精英領導體制 – Meritocracy

一瞬間 – In an instant, in the twinkling of an eye

攢錢 – Save money

拼命 – Exert one’s utmost, desperately

修養 – Self-cultivation; accomplishment, mastery, taste (in art etc.)

替代 – Substitute for, replace, supersede

解脫 – Escape, release, free oneself, extricate oneself

貧困 – Poverty

熬夜 – Burn the midnight oil, pull an all-nighter

珍視 – To value, to treasure, to prize
Appendix K: kàn bú jiàn de ài 看不见的爱

看不见的爱

夏季的一天，天色很好，我决定出去走走。在一片空地上，我看见一个 10 岁左右的男孩和一位妇女。那孩子正用一只做得很粗糙的弹弓（slingshot）射一只立在地上，离他七八米远近的玻璃瓶。

那孩子有时能把弹丸打偏一米，而且忽高忽低。我便站在他身后不远处，看他练习，因为我还没有见过打弹弓这么差的孩子。那位妇女坐在草地上，从一堆石子中捡起一颗，轻轻递到孩子手中，安详地微笑着。那孩子一颗颗接过来，一颗颗打出去，当然，他都浪费掉了。从那妇女的眼神可以看出，她是那孩子的母亲。

那孩子很认真，屏住气，很久才打出一弹。但我站在旁边都可以看出这一弹一定打不中，可是他没有罢手的意思。

我走上前去，对那位母亲说：“让我教他怎么打好吗？”男孩停住了，但还是看着瓶子的方向。母亲对我笑了一笑，说：“谢谢，不用！”她顿了一下，望着孩子悄悄对我说：“他看不见。”我怔住了。半晌，我喃喃地说：“噢……对不起，但为什么？”“别的孩子都这么玩，不是吗？”“呃……”我说，“可是他……怎么能打中呢？”“我告诉他，他总会打中的。”母亲平静地说，“关键是他做了没有。”我沉默了。过了很久，男孩的频率逐渐慢了下来，他已经累了。母亲并没有说什么，还是很安详地捡石子，微笑着，只是递石子的节奏也慢下来了。我慢慢发现，这孩子打得很有点规律，他射出一弹，向一边移一点，再射一弹，再移一点，然后再慢慢地反方向移回来。他只知道大致的方向啊！

夜风轻轻袭来，天幕上已经有了疏朗的星星，石子崩在地上“砰砰”声仍在单调地重复着。对于那孩子来说，黑夜与白天并没有什么区别。

“看来今天他打不中了。”我想。犹豫了一下，我对他们说声再见，便转身向回走去。

走出不远，突然身后传来一声清脆的瓶子碎裂声，随即是划破夜空的，夸张的令人心碎的母子的欢呼声……
看不見的愛

夏季的一天，天色很好，我決定出去走走。在一片空地上，我看見一個 10 歲左右的男孩和一位婦女。那孩子正用一支做得很粗糙的彈弓(slingshot)射一支立在地上，離他七八米遠近的玻璃瓶。

那孩子有時能把彈丸打偏一米，而且忽高忽低。我便站在他身後不遠處，看他練習，因為我還沒有見過打彈弓這麼差的孩子。那位婦女坐在草地上，從一堆石子中撿起一顆，輕輕遞到孩子手中，安詳地微笑著。那孩子一顆顆接過來，一顆顆打出去，當然，他都浪費掉了。從那婦女的眼神可以看出，她是那孩子的母親。

那孩子很認真，屏住氣，很久才打出一彈。但我站在旁邊都可以看出這一彈一定打不中，可是他沒有罷手的意思。

我走上前去，對那位母親說：“讓我教他怎麼打好嗎？”男孩停住了，但還是看著瓶子的方向。母親對我笑了一笑，說：“謝謝，不用！”她頓了一下，望着孩子悄悄對我說：“他看不見。”我怔住了。半晌，我喃喃地說：“噢……對不起，但為什麼？”“別的孩子都這麼玩，不是嗎？”“呃……”母親對我說，“可是他……怎麼能打中呢？”“我告訴他，他總會打中的。”母親平靜地說，“關鍵是他做了沒有。”我沉默了。過了很久，男孩的頻率逐漸慢了下來，他已經累了。母親並沒有說什麼，還是很安詳地撿石子，微笑著，只是遞石子的節奏也慢下來了。我慢慢發現，這孩子打得很有規律，他射出一彈，向一邊移一點，再射一彈，再移一點，然後再慢慢地反方向移回來。他只知道大致的方向啊！

夜風輕輕襲來，天幕上已有了疏朗的星星，石子崩在地上“砰砰”聲仍在單調地重複著。對於那孩子來說，黑夜與白天並沒有什麼區別。

“看來今天他打不中了。”我想。猶豫了一下，我對他們說聲再見，便轉身向回走去。

走出不遠，突然身後傳來一聲清脆的瓶子碎裂聲，隨即是劃破夜空的，誇張的令人心碎的母子的歡呼聲……
Appendix L: kàn bú jiàn de ài 看不见的爱 Questions

1. Why was the lady sitting on the ground?

2. How does the author first describe the lady’s smile?

3. How did the author react when the lady told him why the child kept missing the target?

4. After the author had watched for a while, how did the child’s shooting change?

5. What did the author hear after he left?
Appendix M: kàn bú jiàn de ài 看不见的爱 Vocabulary List with Example

射 shè shoot, fire e.g. an arrow, a bullet, a ball
古时候中国人射箭打猎(hunt)。古時候中國人射箭打獵(hunt)。

捡 jiǎn pick up, gather, collect
请把掉到地上的纸捡起来。請把掉到地上的紙撿起來。

安详 ān xiáng composed, serene, peaceful
老人安详地睡着了。老人安詳地睡著了。

怔住 zhèng zhù stare blankly, freeze
他的行为太过分了(beyond the pale), 周围的人(people around him)都怔住了。他的行為太過分了(beyond the pale), 周圍的人(people around him)都怔住了。

半晌 bàn shǎng half of a day; quite a while, a long time, a short while (meaning depends on context)
我没有复习功课, 老师问我问题时, 过了半晌我才想起来该怎样回答。
我沒有復習功課, 老師問我問題時, 過了半晌我才想起來該怎樣回答。

频率 pín lǜ frequency
最近他给女朋友打电话的频率越来越高。最近他給女朋友打電話的頻率越來越高。

逐渐 zhú jiàn gradually, little by little
和女朋友分手之后, 过了一段时间, 他心情逐渐好起来了。
和女朋友分手之後, 過了一段時間, 他心情逐漸好起來了。

节奏 jiézòu rhythm (of music or figuratively)
我很喜欢这首歌的节奏, 很快。我很喜歡這首歌的節奏, 很快。

疏朗 shū láng thinly scattered, sparse (and therefore clear or obvious, i.e. stars)
夜空中有疏朗的几点星光。夜空中有疏朗的幾點星光。

崩 bēng to hit/strike somebody or something
我踢的足球崩坏了他的眼镜。我踢的足球崩壞了他的眼鏡。

犹豫 yóu yù hesitate
你别犹豫了, 现在就买那台好的电视吧。你別猶豫了, 現在就買那台好的電視吧。
清脆  清脆  qīng cuì  clear and crisp, clear and melodious (like of singing)
她的歌声很清脆，很好听。她的歌声很清脆，很好听。

碎裂   碎裂  suì liè  smash to pieces, break, crack
眼镜掉在车轮下，压得碎裂了。眼镜掉在车轮下，压得碎裂了。
Appendix N: kàn bú jiàn de ài 看不见的爱 Vocabulary List

<table>
<thead>
<tr>
<th>Chinese</th>
<th>Pinyin</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>射/射</td>
<td>shè</td>
<td>shoot, fire e.g. an arrow, a bullet, a ball</td>
</tr>
<tr>
<td>捡/撿</td>
<td>jiǎn</td>
<td>pick up, gather, collect</td>
</tr>
<tr>
<td>安详/安詳</td>
<td>ān xiáng</td>
<td>composed, serene, peaceful</td>
</tr>
<tr>
<td>怔住/怔住</td>
<td>zhèng zhù</td>
<td>stare blankly, freeze</td>
</tr>
<tr>
<td>半晌/半晌</td>
<td>bàn shǎng</td>
<td>half of a day; quite a while, a long time, a short while (meaning depends on context)</td>
</tr>
<tr>
<td>频率/頻率</td>
<td>pín lǜ</td>
<td>frequency</td>
</tr>
<tr>
<td>逐渐/逐漸</td>
<td>zhú jiàn</td>
<td>gradually, little by little</td>
</tr>
<tr>
<td>节奏/節奏</td>
<td>jié zòu</td>
<td>rhythm (of music or figuratively)</td>
</tr>
<tr>
<td>疏朗/疏朗</td>
<td>shū lǎng</td>
<td>thinly scattered, sparse (and therefore clear or obvious, i.e. stars)</td>
</tr>
<tr>
<td>崩/崩</td>
<td>bēng</td>
<td>to hit/strike somebody or something</td>
</tr>
<tr>
<td>犹豫/猶豫</td>
<td>yóu yù</td>
<td>hesitate</td>
</tr>
<tr>
<td>清脆/清脆</td>
<td>qīng cuì</td>
<td>clear and crisp, clear and melodious (like of singing)</td>
</tr>
<tr>
<td>碎裂/碎裂</td>
<td>suì liè</td>
<td>smash to pieces, break, crack</td>
</tr>
</tbody>
</table>
Appendix O: Directions for the Test

Group 1

Please wait quietly for five minutes – you may work on homework. After five minutes, you will be given a passage in Chinese to read. You will have twelve minutes to read. When the time is up, the passages will be collected, and you will be given questions. You will have five minutes to answer these questions in English.

Group 2

You will be given a vocabulary list with example sentences. You will have five minutes to study the words and examples. When time is up the vocabulary lists will be collected, and you will be given a passage in Chinese to read. You will have twelve minutes to read. When the time is up, the passages will be collected, and you will be given questions. You will have five minutes to answer these questions in English.

Group 3

You will be given a vocabulary list. You will have five minutes to study the words. When time is up the vocabulary lists will be collected, and you will be given a passage in Chinese to read. You will have twelve minutes to read. When the time is up, the passages will be collected, and you will be given questions. You will have five minutes to answer these questions in English.

Group 4

You will be given a vocabulary list with example sentences. You will have five minutes to study the words and examples. After five minutes you will be given a passage in Chinese to read. You will have twelve minutes to read. You may refer to the vocabulary list as you read. When the time is up, the passages will be collected, and you will be given questions. You will have five minutes to answer these questions in English.

Group 5

You will be given a vocabulary list. You will have five minutes to study the words. After five minutes you will be given a passage in Chinese to read. You will have twelve minutes to read. You may refer to the vocabulary list as you read. When the time is up, the passages and vocabulary lists will be collected, and you will be given questions. You will have five minutes to answer these questions in English.
Group 6

Please wait quietly for five minutes – you may work on homework. After five minutes, you will be given a vocabulary list with example sentences and a passage in Chinese to read. You will have twelve minutes to read. You may refer to the vocabulary list as you read. When the time is up, the passages and vocabulary lists will be collected, and you will be given questions. You will have five minutes to answer these questions in English.

Group 7

Please wait quietly for five minutes – you may work on homework. After five minutes, you will be given a vocabulary list and a passage in Chinese to read. You will have twelve minutes to read. You may refer to the vocabulary list as you read. When the time is up, the passages and vocabulary lists will be collected, and you will be given questions. You will have five minutes to answer these questions in English.