CHAPTER 1

INTRODUCTION

1.1 Background of the Study

In the 21st century, people live in a knowledge-exploded world. New information is flooded everywhere in newspaper, magazines, and on the Internet. To obtain and integrate the tremendous amount of information, the reading ability is essential. Grabe (2009) points out that to achieve success in the world, reading ability is the key. Moreover, for the purpose of retaining the newly-obtained knowledge, readers should not only be able to understand the reading material in the word- and sentence- level, but also efficiently identify the main idea of the text.

Furthermore, with the coming of English as an international language, most information is printed in English. The first-handed news and information in the world are mostly written in English. The ability to read the English texts has become a requirement for any citizen in the globe. In educational settings, the emphasis on English reading ability is exceptionally apparent. In colleges around the world, students are often required to read textbooks written in English. English reading proficiency, thus, has also been essential for students to succeed in college.

1.2 Statement of the Problem

Being able to read English texts also cannot be overemphasized in Taiwan. College students in Taiwan are commonly required to read textbooks written in English by their content-area (e.g. Western Civilization, Psychology of Learning) course instructors. However, students, in general, have difficulty comprehending the text material because of the amount of unknown words and their failure in getting the gist. Difficulty in comprehending the text material has been hindering the college students' growth in their professional knowledge.

Literature on reading research has clearly demonstrated that the ability to comprehend

and memorize the text content is through getting the main idea of the text (Kintsch, 1998). Previous studies that focused on remedial instruction for reading difficulties have shown that summary-writing can effectively help learners identify main ideas and thus gain better comprehension (Caccamis & Snyder, 2005; Friend, 2001; Oded & Walters, 2001; Radmacher & Latosi-Sawin, 1995). In Taiwan, content-area instructors usually assume that students learn through reading and hence tend not to provide additional help to facilitate students' comprehension of the reading materials. Most college students thus don't receive specific reading training. When thinking about ways to help the learners, the researcher found abundant evidence in support of using summary-writing to boost up reading comprehension. In addition, the incorporation of CALL in EFL classroom has been receiving a lot of attention. The combination of the two came to the researchers' attention in the design of this study.

In recent years, the development of computer technology has offered teachers another way to teach languages. The thriving of Computer Assisted Language Learning (CALL) since 1980s has been considered beneficial in several aspects of language learning. Research has pointed out that generally CALL encourages autonomous learning and engages learners more actively in language learning. Most studies have shown that learners in general hold a positive attitude toward the implementation of CALL in language learning (Ayres, 2002; Dodigovic, 2002; Son, 2007), and studies focusing on the effects of CALL on learners' development of specific skills also confirmed the effectiveness of CALL (Sullivan and Pratt ,1996; Braine, 1997; Sun, 2007) on students' writing development.

Seeing the great promise of CALL and recognizing the benefits students can get through summary-writing training in boosting their reading comprehension level, the researcher conducted the research using WriteToLearn, an online summary-writing software, to help learners in Taiwan. The effects of the WriteToLearn summary-writing training program on L1 learners' reading and writing performance were investigated by research such as Wade-Stein & Kintsch (2004), Franzke, Kintsch, Caccamise, Johnson, & Dooley (2005), Caccamise, Franzke,

Eckhoff, Kintsch & Kintsch (2007, as cited in Landauer et al. 2009). Positive findings were reported. However, little research has been done to investigate the effects of WriteToLearn summary-writing training program in EFL setting such as Taiwan. To examine the possible benefits the WriteToLearn program can provide to Taiwanese EFL college learners, the present study is needed. To be specific, the present study aimed at exploring the effects of an on-line summary-writing training through WriteToLearn on Taiwanese EFL college learners' improvement in reading comprehension, vocabulary knowledge and summary-writing.

1.3 The Present Study

The present study aimed at investigating the effects of WriteToLearn summary-writing training on EFL college students' language development in different aspects including reading comprehension, vocabulary learning and summary-writing. The researcher selected the Taiwanese English-major college students as the participants. The English majors in the research site took a one-year course entitled "Guided Reading". This was the only required course they took within the 4-year study period in college that focused on developing reading strategies such as clarifying the purpose of reading, identifying the main idea, skimming and scanning, predicting, inferencing, and guessing from the context. This was also the only required course that provided students opportunities to focus on reading non-fiction articles (such as newspaper and magazine articles). The WriteToLearn summary-writing training offered through this course mainly intended to provide the learners opportunities to practice summary-writing so as to improve their reading comprehension. Moreover, the automated feedback generated by the WriteToLearn provided the learners instant responses for them to revise their summaries. It was hoped that the findings would shed light on the curriculum design of this course.

1.4 Research Questions

The study was designed to examine the effects of summary-writing training of

WriteToLearn on facilitating the EFL college learners' reading comprehension, vocabulary knowledge and summary-writing development. Moreover, to examine to what extent the EFL learners at two different ability levels (high versus low) might benefit from WriteToLearn, the EFL learners were divided to the high-level or low-level reading groups. The present study aimed at answering the following research questions:

1. Do the two groups of EFL learners (high and low levels) respectively improve their reading comprehension upon the completion of the *WriteToLearn* program? Are there any significant differences between the two groups?
2. Do the two groups of EFL learners (high and low levels) respectively improve their vocabulary knowledge upon the completion of the *WriteToLearn* program? Are there any significant differences between the two groups?
3. Do the two groups of participants (high-level and low-level) respectively make progresses in their summary writing during the 14-week training session using WriteToLearn? Is there any significant difference between the two groups in their progresses of summary writing?

1.5 Definition of Terms

Reading Comprehension

Reading comprehension refers to the ability to understand the information provided in the text. Kintch and van Dijk (1978) proposed that comprehension of a text consists of two levels, microstructure and macrostructure (as cited in Hudson, 2007). The microstructure represents the local information in a text, whereas the macrostructure represents the global structure of a text. In this study, a reading test was used to measure learners' reading comprehension. The test contained expository texts. Multiple-choice questions were asked to test both micro-level and macro-level comprehension.

Vocabulary Knowledge

According to Nation (2001), vocabulary knowledge can be divided into three dimensions:

word form, word meaning and word usage. This study measured the EFL learners' vocabulary knowledge using Vocabulary Knowledge Test, which was adopted from Vocabulary Knowledge Scale (VKS) (Paribakht & Wesche, 1993). The Vocabulary Knowledge Test was used to measure the test-taker's receptive knowledge of recognizing word form and word meaning. In addition, it also measures their productive knowledge of the word's grammatical functions by asking the test-takers to make sentences based on the words.

Summary-writing

Summary-writing consists of the ability to distinguish the important content from irrelevant details in a piece of reading, to combine those important content into a summary, and to use one's own words when writing the summary rather than copying words from the original article. In this study, the students' summary-writing performance was examined in 4 aspects which were the major feedback generated by the WriteToLearn program. The four aspects included copying behavior, spelling mistakes, repeated content and unimportant content, all of which should be avoided in writing a summary. Copying behavior refers to the act of copying the words directly from the original passage. Spelling mistakes refer to the wrong spelling of words in a summary. Repeated content are the unnecessary repetitive sentences which convey the same information. Unimportant content are the irrelevant or unimportant details. In this study, the percentages of the 4 aspects were automatically reported by the WriteToLearn program. The lower the percentage is, the better the summary-writing performance is.

1.6 Significance of the Study

The WriteToLearn summary-writing program provides the learners multiple opportunities to write summaries and revise them according to the instant feedback generated by the program. Several studies (e.g. Franzke et al., 2005; Wade-Stein and Kintsch, 2004) have claimed the effects of this program on L1 learners' development. Research is needed to further explore the usefulness of the program in the EFL setting. This study is one of the first attempting to use an

on-line summary-writing program to help Taiwanese EFL learners develop reading comprehension. An investigation on the WriteToLearn program's effects on EFL college learners' reading comprehension development is important as it can inform English educators about EFL learners' needs and shed light on directing the English curriculum development in Taiwan.

CHAPTER 2

REVIEW OF THE LITERATURE

2.1 The Importance of Reading

In our daily life, we conduct reading activities all the time. Adults read newspaper, professional books, and magazines for getting the up-to-date information, while teenagers and children read textbooks for schooling. Why is reading so important? In the present world which is exploded with tremendous amount of information, knowledge is power. Human beings conserve their wisdom through printed words. As such, people need to acquire knowledge from the print. Nowadays, new information often appears on the Web. As Grabe (2009) rightly pointed out that electronic communication, which has become indispensable throughout the global village, has made efficient reading skills even more crucial than ever, for people need to digest the overwhelmingly large amount of information around them. Reading has been undoubtedly an essential ability in the modern society.

The importance of reading efficiently is not limited to the L1 environment. With English having become a global language, a large percentage of people around the world learn English as an L2 for different purposes including doing business, gaining professional knowledge, traveling, or entertaining. The rise of English as a global language has also created a profound impact on the educational system. College education worldwide requires students to gain professional knowledge through reading textbooks written in English. Thus, being able to read the English texts has become a decisive ability in college education. Given the indispensable need of reading English texts, college education should help students develop the ability of comprehending reading materials written in English so that they are capable of obtaining professional knowledge in their subject matter.

2.2 Reading Comprehension

In account for the reading process in English as an L1, researchers have had various

views, which can generally be categorized into two approaches—bottom-up approach and top-down approach (Hudson, 2007; Kintsch, 2005). Bottom-up approaches assume that readers decode individual words from letters and subsequently proceed to the sentence level to construct meanings. When reading, readers proceed simply from left to right and passively receive information from the text (Gough, 1972, as cited in Hudson, 2007). This approach assumes that readers' background knowledge and personal experience have little effect on the processing of the text. Readers comprehend the text merely through the information provided by the text. Good readers can automatically match the visual code with their mental lexicon, thus reducing the working memory needed for decoding and leaving more room for text comprehending. Successful reading comprehension is then achieved (LaBerge & Samuels, 1974, as cited in Hudson, 2007). From this perspective, word identification and the speed of text processing are the keys to successful reading (McConkie & Zola, 1980).

Top-down approaches, in contrast, suggest that background knowledge has a role to play in the reading process. With the previous knowledge stored in memory, the readers already have had a working hypothesis for the content of the text before reading. During reading, the readers modify their hypothesis based on the new information provided in the text. In other words, the readers use their background knowledge to construct the meanings of the text that is "personally and contextually sensible" (Hudson, 2007, p.34). As such, reading is an active process in which readers tend to select meaningful contents based on their background knowledge on the topic and construct sensible text meanings for themselves.

A more balanced view has been proposed in recent years and termed interactive approaches. Interactive approaches claim that reading involves the readers' personal background knowledge and the information provided in the text. Moreover, the interactive view acknowledges that in the reading process, many variables are interacting with one another (Kintsch, 2005). The interactive approaches mainly focus on three kinds of interactions during the reading process: 1) the lower-level skills implicated automatically by the readers; 2) the

interaction between readers' background knowledge and the text; 3) the interaction of social, contextual and political variables with the readers' interpretation of the text.

The three kinds of interactions are described as follows. The first kind of interaction is illustrated by Rayner and Pollatsek (1989). They stresses that the emphasis of reading comprehension is not on the product of reading, but the process. They assert that to understand reading, people should understand the process through which the readers construct their mental structures. These lower-level cognitive processes, such as eye-movement and decoding words, have more direct influence on reading than the context. Another kind of interaction is between readers' background knowledge and the text. Smith (1994) states that language contains two structures—surface structure (i.e. the surface characteristics of the sentences) and deep structure (i.e. the meaning that the readers obtain from the sentences). The reading process should focus on the interaction between the two structures, that is, between the writer's original intention and the reader's interpretation. The third kind of interaction emphasizes the importance of social context. That is, readers take into consideration the context of reading, such as the purpose of reading, or the genre of the text. For example, whether a reader is supposed to read critically or just for entertainment would affect the reading process. A reader whose purpose is to read critically would read more carefully and would probably read the text more than once and attempt to reflect on the content of the text from different points of view. Those who read for entertainment may only read through the text in a rapid pace with little reflection upon the text. Similarly, the genre of the text affects the readers' comprehension and memorization of the text. For instance, narratives are more easily to be retained than expositories.

In sum, the interactive perspective of reading involves not only the text itself and the background knowledge of the readers, but also the context in which reading occurs. Nowadays, few researchers take a pure bottom-up perspective of reading, nor does anyone support a strong top-down view. The most current pedagogical approaches are interactive approaches, which

acknowledge a significant amount of communication between the views proposed by the two extreme perspectives (Hudson, 2007). The interactive view of reading process is well represented by a model called Construction-Integration Model (the CI model), which is proposed by Kintsch (1988). The CI model combines the bottom-up processing and the top-down processing in reading. The term "Construction" refers to the readers' processing of text information through word recognition or syntactic analysis. The term "Integration" refers to the process in which the readers integrate the incoming information into a coherent statement of the text by identifying the repetitive information of the prior content or the strong association with the most active information provided already. That is, the main idea of the text is identified through the process in which the readers integrate the text information by extracting the most repetitively-mentioned information from the text. Clearly, construction occurs at a more local level and integration occurs at a more global level. The following section provides a more clear description about the two levels.

2.3 Macrostructure and Microstructure

According to van Dijk and Kintsch (1978), a reader comprehends a text at two levels, macrostructure level and microstructure level. The microstructure level is the local level of the text which provides the structure of individual sentences and the relations among them. On the other hand, the macrostructure of a text is at the global level, representing the gist of a text. In other words, the microstructure represents the readers' understanding of the regional knowledge of the text, while the macrostructure represents the main idea of the text, which is acquired through finding the coherence of several representative micro-level concepts.

Van Dijk and Kintsch stated that microstructure is the surface structure of the text, representing the semantic meaning of individual words and demonstrating the relations among them. In a text, the microstructure provides the basis for the general information of the text. During the process of reading, readers use macrorules to extract the macrostructure of the text from the microstructure level. There are three macrorules: deletion, generalization and

construction (Kintsch & van Dijk, 1978). Deletion is the process of removing the unimportant propositions that do not affect the interpretation of the following propositions. Generalization is the process in which readers use a general term to represent a series of micropropositions. For example, the sentences "Mary was drawing a picture. Sally was jumping rope and Daniel was building something with Lego blocks" may be generalized into "The children were playing" by the reader (Niska, 1999). Lastly, construction refers to the readers' constructing new macro-propositions for a sequence of micropropositions in accordance with general knowledge(see page 9). For example, the sentence "John went to the station. He bought a ticket, started running when he saw what time it was and was forced to conclude that his watch was wrong when he reached the platform." can be condensed into the macrostructure "John missed the train," since the reader construct the words "train" and "miss" from the original microstructure on the basis of their general knowledge (Niska, 1999). By using the three macrorules, readers extract the macrostructure of the text from the microstructure, thus processing the text in a deeper level. Finally, the macrostructure, or the gist of a text, can be stored in the readers' memorization more easily than the microstructure. In sum, the microstructure of a text provides the basis of general knowledge for the readers. In the reading process, the readers apply macrorules to extract the macrostructure, or the gist of a text, from the microstructure. Comprehension of a text is thus achieved.

While recognizing the importance of the microstructure in reading comprehension, Kintsch claims that "for comprehension and memory, the gist of a text—expressed formally by the macrostructure—is usually what matters most" (p.67, as cited in Koda, 2004). To compare the contributions of knowledge of macrostructure and knowledge of microstructure, Britt and Sommer (2004) examined the effect of macrostructure summarizing of prior text on subsequent text comprehension and integration. Their results showed that the students who had done a macro-summary task of the first text performed better than students who had not done such task in the subsequent task which required them to integrate the cross-text information.

Moreover, the students required to answer macro-level questions after reading the first text performed better in the content recall of both texts than those who were treated micro-level questions. In sum, their findings imply that macrostructure knowledge made greater contribution to successful text comprehension and cross-texts knowledge integration than microstructure knowledge.

2.4 Summary-writing Training and Reading Comprehension

Along the same line of research, numerous studies have found that summary-writing training facilitates the reading comprehension performance of students with learning disabilities (Gajria & Salvia, 1992; Jitendra, Cole, Hoppes & Wilson, 1998; Jitendra, Hoppes & Xin, 2000; Rogevich & Perin, 2008). For instance, in Gajria and Salvia's study (1992), they randomly assigned thirty 6-9 graders with learning disabilities into experimental and control groups. The experimental group received the training of explicit summarization skills— (a) superordination, (b) deletion of redundant information, (c) selection, (d) invention, and (e) deletion of unimportant information, whereas the control group did not. The experimental group was taught the five skills on the basis of the principles of direct instruction: explicit explanation of the rules, modeling the strategy, guided practice in controlled materials, monitoring with corrective feedback, and independent practice. The experimental group received the direct instruction on each of the five skills, and practiced each of these rules until they had qualified performance. Finally, the participants used all five rules to construct oral summaries of expository passages until they reached mastery of the five summarization rules. Six expository passages, each of which contained 10 multiple-choice comprehension questions, were used for pretest and posttest. The results showed that the students who received the summary-skill training improved significantly in their performance in reading comprehension as measured by the posttest.

Summary-training not only help learners with reading disabilities improve reading comprehension but also facilitate average learners' reading comprehension (Friend, 2001;

Oded & Walters, 2001; Radmacher & Latosi-Sawin, 1995). For example, Radmacher and Latosi-Sawin (1995) used summary-writing to improve the text comprehension of 16 students in a psychology course. The students discussed the quality of four summaries of the same passage and then rank ordered them with 1 representing "needs least revision" and 4 representing "needs most revision". Through the discussion led by the instructor, they shared their analysis of the four summaries and finally discovered the criteria for a successful summary. Afterwards, every week the experimental group were required to write a summary for a section of the psychology course content and bring their summaries to class for discussion. In class, they wrote the key points in their summaries on the blackboard and discussed which key points were essential to understanding the text content. In contrast, the control group used the same course material in class but took text-comprehension exams. Before the treatment, the two groups did not have significant difference in their performance on text-comprehension pretest. After the treatment, the text-content comprehension test was used for final evaluation of learning. Results showed that the final text-comprehension mean score of students treated with summary-writing is 8% higher than that of the students who had not been treated with summary-writing. Also, in the course evaluation, two-thirds of the students in summary-writing class responded that they found themselves improve in understanding and retention of the text material.

Another study by Friend (2001) compared the effects of two different reading strategies on college freshman's summary performance: argument repetition and generalization. Argument repetition, regarded as a microstructure processing, focused on building coherence in the text. Instruction in argument repetition guided the students to pay attention to the repeated references in the passage, thus students could identify what is considered the most important by the author. Generalization, regarded as macrostructure processing, was a process through which the reader identifies the hierarchical structure of the ideas and thus abstract the gist out of a passage. By generalization, readers built a gist on the basis of the similarity among

a sequence of microstructure knowledge. Based on a sequence of the microstructure sentences such as “Mary was drawing a picture. Sally was jumping rope and Daniel was building something with Lego blocks,” a reader may generalize the gist “The children were playing.” In Friend’s study (2001), one hundred and forty-seven freshmen who were identified as poor writers were assigned to three groups (two experimental versus one control). The two experimental groups were taught to use argument repetition strategy and generalization strategy respectively in summary-writing. The participants in the control group were taught to do self-reflection in which they did their best to understand the gist of the text by reflecting on what is the most important in the text by themselves when writing a summary. Analyses of the summaries demonstrated that the two experimental groups performed significantly better than the control group in judging the importance of text content. The result indicated that summary-writing strategy training helped the students better identify the main idea of a passage. Moreover, Friend also found that the generalization group performed significantly better than argument repetition group in writing a thesis statement. Since the text-processing theory of van Dijk and Kintsch (1983) categorizes generalization as a macro-level strategy and argument repetition as a micro-level strategy, the result implies that macro-level processing is a more effective strategy than micro-level processing in identifying the thesis statement of a text.

2.5 Vocabulary Acquisition and L2 Reading Comprehension

Numerous studies have confirmed that vocabulary knowledge is a critical component of reading comprehension (Joshi, 2005; Landi, 2010; Lervag & Aukrust, 2010), which is especially true in L2 reading. Lervag and Aukrust (2010) claim that vocabulary knowledge is a good predictor for L2 reading comprehension growth. Along the same line of research, Just and Carpenter (1985) report that “across numerous studies, the correlations between vocabulary knowledge and comprehension ranges from 0.66 to 0.75” (cited in Joshi, 2005, p.210), which indicate strong association between vocabulary knowledge and reading comprehension ability. Koda (2004) also argues that in comparison with other factors, vocabulary knowledge makes

the greatest contribution to reading comprehension. The results of these studies all indicate that vocabulary is an important component of reading comprehension, and L2 readers with better vocabulary knowledge have better reading comprehension of texts.

Moreover, vocabulary knowledge is not only a crucial component of reading comprehension, but also a "by-product" of reading comprehension development (Koda, 2004). Previous studies show that reading can enhance vocabulary learning (Brown, Waring & Donkaewbua, 2008; Day, Omura & Hiramatsu, 1991; Nagy, 1995; Pitts, White & Krashen, 1989; Rodriguez & Sadoski, 2000). For instance, Day et al. (1991) investigated the effect of story-reading on vocabulary gain. They had the Japanese-speaking college and high-school learners read English short stories silently and then had the learners take the vocabulary test in which the target words were chosen from the stories. The control-group, on the other hand, was not provided with the stories before taking the vocabulary test. The result showed that the students who read the stories scored significantly higher on the vocabulary test than those who did not. Day et al. concluded that Japanese college and high-school EFL learners could acquire previously unknown or unfamiliar vocabulary simply by reading meaning texts.

Another study by Waring and Takaki (2003) investigated how many words L2 learners could learn and retain from reading a graded reader. The researcher examined the effects of word frequency and test type on word acquisition. The participants were fifteen female Japanese university students with lower-intermediate English proficiency. They read *A Little Princess* and took three tests in different formats-- word-form recognition test, multiple-choice recognition test and meaning-translation test. Afterwards, the participants took the first post-test seven to ten days later and the second post-test after one month. Waring and Takaki (2003) found that a small proportion of new words were learned, and half of the words acquired were not retained after three months. They also reported that word frequency indeed had an effect on acquisition. Moreover, different test types did have significant impact on their performance. Though only a small proportion of unknown words were acquired after reading

the graded reader, Waring and Takaki concluded that reading tends to have lasting effect on word learning.

Researchers, in general, believe that reading provides the learners the contextualized information in which the learners pick up relevant semantic knowledge of the words and integrate it in their mental lexicon. Reading provides the learners the context in which the acquisition of vocabulary knowledge can be enhanced (Miller, 1995; Nagy, 1995; Rodriguez & Sadoski, 2000; Webb, 2008). Nagy (1995) claims that context is undoubtedly necessary in vocabulary acquisition. Moreover, he also suggests that in vocabulary acquisition, second-language learners benefit more from context learning than native learners. Nagy stated that this may be because in context learning, L2 learners encounter more unknown words than L1 learners. In all, providing context for L2 learners in vocabulary learning contributes significantly to L2 learners' learning of vocabulary knowledge.

2.6 The effect of CALL on Writing and Summarization

During the past decades, CALL (Computer Assisted Language Learning) has thrived in the field of language learning. CALL is an approach to language learning which applies computer technology to support the "presentation, reinforcement and assessment" of learning material (Davies, 2000, p.90). Computer technology has been increasingly applied in language learning, and a flourishing number of research has been done to investigate the effectiveness of CALL. According to Cobb (2002), computer technology and CALL provide a great number of benefits for the acquisition and application of English language skills, and teaching English with computers can also enhance students' motivation and confidence in using the English language (as cited in Lee, 2007). Loucky (2003) asserts that implementation of CALL for both L1 and ESL or EFL language learning should be introduced to and used by language teachers so that instruction, motivation and learning would be enhanced. For its vast positive effects on English language learning, CALL has been considered a useful approach and has been widely implemented in all levels of language learning.

The section focuses on research on the effect of CALL in writing in general, and on summary writing in particular. Previous research has found that CALL enhances EFL learners' writing performance (Braine, 1997; Sullivan & Pratt, 1996; Sun, 2007). For example, Sullivan and Pratt (1996) examined Spanish ESL students' writing quality in two different classroom settings: a networked computer-assisted classroom and a traditional oral classroom. In the networked computer-assisted classroom, the students gave responses and carried out discussions on peers' writing through an electronic discussion program named InterChange. In the traditional oral classroom, students discussed their compositions and gave oral responses to peers. The result showed that the students in the networked computer-assisted classroom did improve significantly more in their writing quality than those in traditional oral classroom. The researchers argued that students making discussions through the network seemed to have more time to read and reflect before responding, which seemed to be beneficial for ESL student writers. Braine (1997) conducted a similar study comparing ESL students' writing performance in two different contexts: a networked computer class and a traditional oral-discussion class. Throughout the treatments, the students were required to write three compositions, have discussions on one another's compositions and then provide feedback to peers. The networked computer class did their work through the network, whereas the traditional oral-discussion class did their work through face-to-face discussion and written comments. Braine found that the students in networked computer class wrote significantly longer compositions than those in the traditional class. Also, they scored significantly higher than the traditional oral-discussion students both in their first and final compositions. The better writing quality in networked computer class, as Braine suggested, might be because the networked computer environment reduced the ESL learners' anxiety in learning. Moreover, the researcher commented that the networked computer environment reduced the amount of teacher talk, which was typical in traditional classroom. Through the online discussion, interaction among learners increased in a truly communicative situation, which promotes effective language learning. In an EFL setting,

Sun (2007) examined the Taiwanese learners' perceptions of an online Scholarly Writing Template(SWT), which provides students with information guidance of organization moves and a language template. Sun found that students' writing performance improved after the SWT treatment and their writing strategies became more delicate and ingenious. Moreover, students acknowledged the effectiveness of the SWT in terms of helping them improve writing skills. Overall, these studies all suggest that implementation of CALL has a positive effect on ESL and EFL learners' writing quality.

In addition to their reporting of the positive effects CALL has on learners' writing quality, researchers also found that ESL and EFL learners generally hold positive attitude toward CALL (Ayres, 2002; Dodigovic, 2002; Son. J, 2007). Ayres (2002) examined undergraduate ESL learners' attitude toward computer usage in language learning. He surveyed the learners' views about the usefulness of various language-learning software that they had been using in their course. The analysis of the questionnaire showed that 80% of the learners considered the software meeting their learning needs, 77% thought the information provided by the software was useful, and 60% suggested that CALL should be implemented more often in language learning. However, the learners showed preference for classroom-based teaching over computer using. Ayres concluded that most ESL learners perceive CALL as a useful tool in language learning, which enhances, but not replaces the classroom-based instruction (p.249). In addition, Son (2007) also examines the perceptions and attitudes of ESL learners toward Web-based Language Learning (WBLL) in an English language intensive course for overseas students (ELICOS). The results showed that most of the learners consider WBLL a useful tool for learning and are willing to use it outside the classroom. Son concluded that the participants hold a positive attitude toward the web-activities.

In sum, the research findings suggest that ESL learners have positive attitude toward CALL and consider the implementation of CALL effective in facilitating language learning. In addition, summary-writing training indeed enhances students' reading comprehension

performance. Therefore, in the present study, the researcher further explored this field by implementing an online software, WriteToLearn, in the reading, and examined the effects of this on-line training on learners' learning in several aspects. WriteToLearn was used as a supplementary tool for, not a replacement of, the classroom-based instruction. The following paragraphs introduce the WriteToLearn program and review previous research on the use of it.

2.7 The WriteToLearn Program

WriteToLearn is an online Web-based system which provides students expository passages suitable for learners at various levels to write summaries for. The system offers automated feedback on the quality of the summaries. After completing their summaries on line, students can submit their summaries to the system to receive feedback. Then they can revise their summaries 5 more times at most based on the feedback and resubmit to receive new feedback.

The method that WriteToLearn uses to evaluate students' summaries is Latent Semantic Analysis (LSA). LSA analyzes the meanings and relatoinship of words, sentences and articles in a mathematical way, similar to vectors in "a high-dimensional semantic space" (Wade-Stain & Kintsch, 2004, p.337). The hypothetical semantic space was built up by examining the meaning of millions of words which occur in thousands of documents. Words with similar meanings or occurring in similar contexts are located near one another in this space, and so as texts with similar meanings. This semantic space enables LSA to calculate the "semantic distance" between the meanings of two different words or passages. LSA measures the semantic distance between words or passages by using the cosine, which can be seen as an index of correlation. Words that are highly related would have the high cosine with a maximum value of 1, whereas words that are unrelated have the low cosine with a minimum value of 0. When LSA is used in WriteToLearn, the cosine between the content of the original texts and that of a student's summary is calculated. In this way, the content of the students' summaries is evaluated.

According to Wade-Stein and Kintsch (2004), LSA does not calculate the co-occurrence among words, but the semantic relations among them. That is, LSA calculates the similar semantic meaning in different words instead of searching for the repeated occurring of the same word. Thus, LSA is suitable for summary-evaluation, since a summary is not supposed to contain the same words as those in the original text. With LSA, WriteToLearn is able to compute the correlations between the content of the summary and that of its original text, thus giving the learners feedbacks as to how well the summary covers the content of the original text.

The major weakness of LSA is that it only measures the semantic aspect of an essay. It does not evaluate syntax, word order or the flow of arguments (Kintsch, et al., 2000; Foltz, Gilliam & Kendall, 2000). The criticism of LSA asserts that being unable to measure these qualities, LSA cannot evaluate the overall quality of an essay precisely. To examine whether LSA's measure of semantic content could successfully represents the measure of the overall quality of an essay, Foltz et al. (2000) investigated the correlation between the quality of overall writing and the quality of content. The result shows that the correlation is 0.76. The results indicate that the scoring for the content of the essay can accurately account for the overall quality of an essay. Based on the findings, WriteToLearn, which applies LSA to evaluate students' summaries, is able to make fair judgment about the quality of the summaries holistically.

The WriteToLearn online program reflects the quality of learners' summaries by providing learners immediate feedback on their summary-writing performance. Using WriteToLearn, learners can receive scores on the content, length, spelling and relevancy of their summaries immediately after each time they submit the first-draft or revised-draft of their summaries. The immediate, frequent and individualized feedback to learners might not be easily achieved in traditional classroom settings (Landauer, Lochbaum & Dooley, 2009). Moreover, previous research shows that most students enjoy the repeated trials with improving

scores shown on the screen (Landauer, Lochbaum & Dooley, 2009).

In addition to the instant feedback on learners' summaries, researchers also found that the WriteToLearn summary-writing training benefits learners' reading and writing performance. Wade-Stein and Kintsch (2004) had two 6th grade EFL classes write summaries. During the first week of training, half of the students composed summaries and revised them in the WriteToLearn system, which provided instant feedback for their summaries. The other half did the summarization and revision by using a similar interface, which merely displayed the summary-length indicator without content feedback. During the second week of training, the students did the same tasks but with the conditions counterbalanced. The students who had used WriteToLearn in the previous week did the summaries without feedback, while those who had not used WriteToLearn got on the WriteToLearn training program and got instant feedback. Subsequently, the students took turns composing summaries in the feedback or no-feedback condition. The result showed that summaries receiving instant feedback from WriteToLearn scored significantly higher in content by the teacher graders than summaries receiving no feedback. Moreover, students spent more than twice amount of time working on their summaries in the condition where content feedback was provided, compared to students in the no feedback condition. Wade-Stein & Kintsch concluded that receiving content feedback on their summaries indeed help learners write better summaries.

Similar results have been found in Franzke et al. (2005). Franzke and other researchers had two groups of 8th-grade learners write summaries. The experimental group used WriteToLearn while the control group used a common word processor without feedback. Four weeks later, the group using WriteToLearn had a significant better increase in the content score of their summaries than the control group. Moreover, the researchers found significant greater gain in the experimental group not only in skills of condense and abstract but also in organization and style. Finally, researchers found that students using WriteToLearn wrote summaries which contained 50% more relevant content than those who wrote summaries

without WriteToLearn (Caccamise, Franzke, Eckhoff, Kintsch & Kintsch, 2007, as cited in Landauer et al. 2009; Franzke et al., 2005). The above-mentioned studies suggest that WriteToLearn facilitates learners' summary-writing development by providing them with automated feedback.

Based on the previous findings which demonstrated that WriteToLearn summary-writing training enhances learners' growth in writing ability, the present study aimed at investigating the effect of WriteToLearn on Taiwanese college EFL learners' development in reading comprehension ability, vocabulary knowledge growth, and summary-writing ability. This study further explored to what extent the EFL learners at two different ability levels benefit from WriteToLearn.

CHAPTER 3

METHOD

3.1 Participants

The participants in this study were 33 freshman English-majors in a university located in central Taiwan. They were composed of 9 males and 24 females. Except for 2 delay-graduated students and 2 sophomores, all the participants were college freshmen aged 18-20. They all spoke Mandarin Chinese as their L1. Only three of them had short-term study-abroad experiences in English-speaking countries (1-3 months) and none of them had received any summary-writing instruction. In the research site, freshman English majors took Guided Reading as a required course for a year. In the first semester, they learned essential reading strategies (i.e., locating the topic sentence, finding the main ideas) through intensive and extensive reading. In the second semester, they received the summary-writing training. As stated in Chapter 1, the researcher aimed at comparing the effects of the summary-writing training on high-level versus low-level learners' language development. The participants, hence, were grouped into a high-level reading group or a low-level reading group, based on their performance on the final examination they took at the end of the first semester. The examination was considered a direct way to measure students' reading comprehension ability as students were required to judge which statement best represents the main idea of the passage, to give an appropriate title to a passage or to write a short sentence stating the gist of the passage. The range of their final exam scores was between 59 and 97. The researcher used 79 as the cut-off point to divide the participants into two groups, making the two groups' number of participants equal. Minor adjustment of group was made according to the reading-comprehension pretest scores. Finally, seventeen students who had scored at or above 79 belonged to the high-level group; sixteen students who had scored below 79 belonged to the low-level group. To ensure there was significant difference between the two groups, the researcher conducted an independent t-test at a significance level of .05. The result confirmed

that a significant difference did exist between the two groups in their reading ability before they received any treatment (t= 8.050, p= .000).

3.2 Preparation for and Implementation of the WriteToLearn Program

In the fall semester of 2010, the participants used an English-learning textbook that mainly focused on reading comprehension strategies. The strategies the participants were taught in class included predicting the content of the article by reading the title, headings and pictures, figuring out the main ideas of each paragraph through topic sentences, supporting details and concluding sentences, using graphic organizers to sketch how the main ideas are related, and using context clues to identify the meanings of unknown words.

In the spring semester of 2011, the participants received summary-writing instructions for two hours in total in the first week. They first reviewed the comprehension strategies and learned how to write a summary. They were guided to write summaries step by step through getting the main idea, leaving out the unimportant details, and using their own words to write a summary instead of copying the original ones from the text. They practiced writing a summary for a short text and then discussed the summaries with their classmates.

After the summary-writing introductory session, the students were introduced to the WriteToLearn training system. In the third week of the spring semester, the participants were taken to a computer lab to familiarize themselves with the interface of the on-line program, learn the features of the program and practice operating the WriteToLearn summary-writing training system. They read one passage provided by WriteToLearn, wrote a summary for it, practiced submitting the summary to the online system, and then received an automated feedback from the system. To explain the entire process of summary-writing, the researcher used the following figures in the computer lab. Figure 1 showed that after reading a passage, the students could start composing summaries in the window. An expected summary length (a range of numbers of words used) would be given by the system under the passage title to give students a clue about an appropriate length of the summary. While reading the passage,

students could highlight any sections and click on the green triangle in the text tools and hear the section read outloud by a native speaker. The red square could be clicked on to stop the speech. The bidirectional arrows were of no practical use for the students because they gave a Spanish translation equivalent of the English word. The most useful tool was the dictionary function. Students could highlight words and click on the dictionary icon and get to know the meanings of the words explained in English. After finishing a summary, they could click on the button "Check Spelling" to have the system check the word spelling for them. After the spelling-check, they clicked on "Get Feedback" to submit the summary to the system and get feedback.

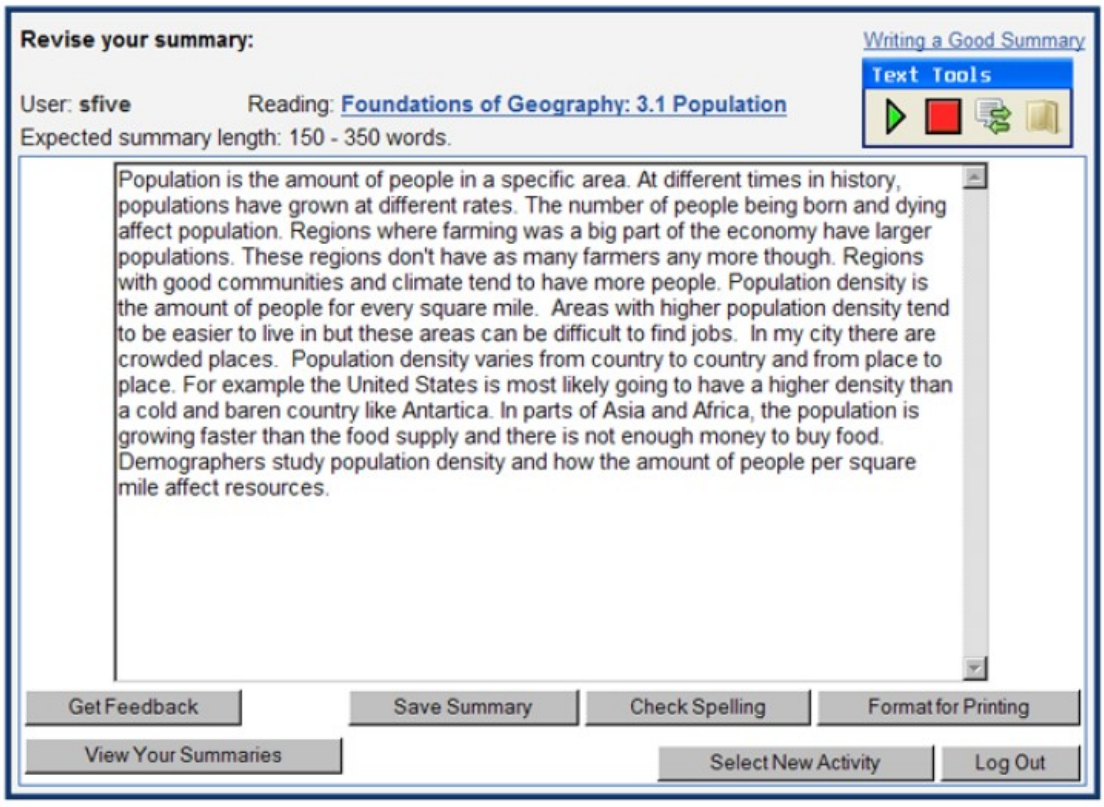

Figure 1. Sample students screen: Summary writing window.

The feedback provided by the WriteToLearn system was shown in a format using scoreboards as illustrated in Figure 2. There were three types of feedback. In the middle were the scoreboards for each section coverage, reflecting how thoroughly the students' summary covered the important content of the major sections in the original passage. The feedback was

shown by the blue bars. The goal of the students was to get the blue bar to move into the green Excellent zone. On the upper right corner was a vertical blue bar showing the feedback on the length of the summary as a whole for the entire passage. As can be seen, neither "too short" nor "too long" was considered good. Students were advised that a good summary stated the main points at an appropriate length. Detailed or irrelevant information should not be included.

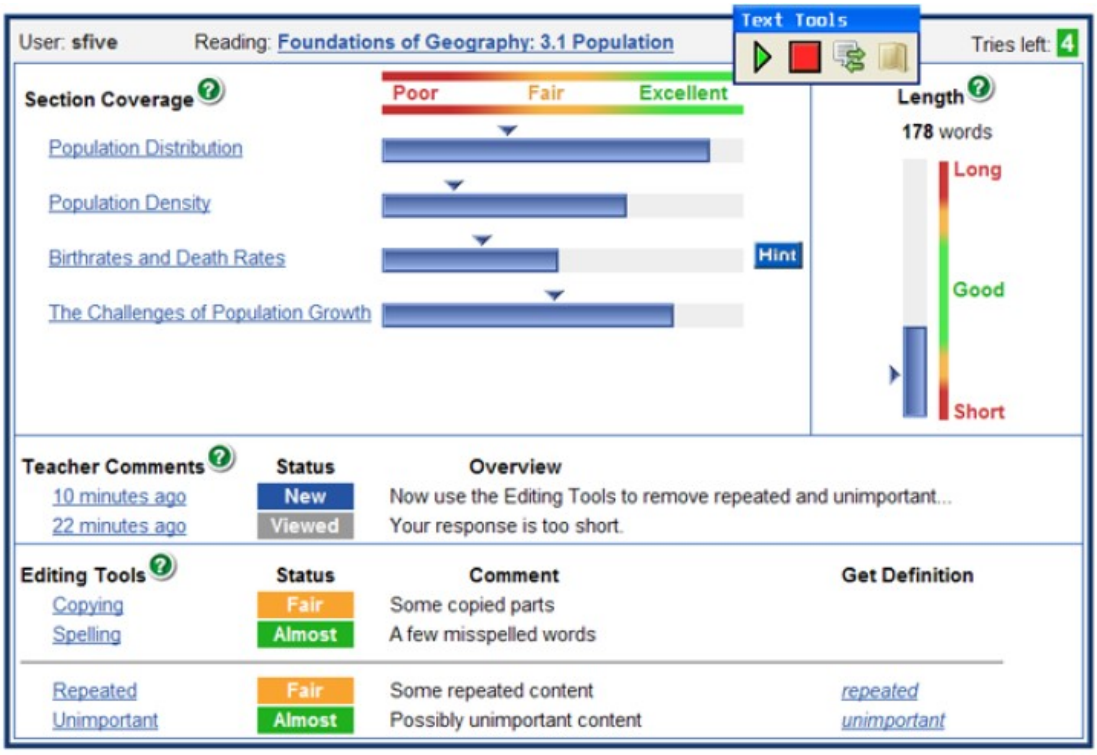

Figure 2. Sample student screen: Graphical display of automated feedback.

Finally, at the bottom of the automated feedback screen, the system provided four aspects of feedback: copying behavior, spelling mistakes, repeated content and unimportant content, all of which should be avoided. First of all, "Copying" gave an estimate of words that were taken directly from the original passage. A "fair" showed that some parts of the summary were taken directly from the original texts without any changes and implied that the student had not yet fully developed good summarization skill to restate the main ideas using his/her own words. Secondly, "Spelling" pointed out the spelling mistakes in the summary. An "almost" showed there were only a few misspelled words in the summary draft. Thirdly,

"Repeated" gave an estimate of sentences that were similar in meaning. A 'fair" indicated some redundant information in the summary draft. Fourthly, "Unimportant" reflected the amount of content in the summary that contained irrelevant or unimportant details. An "almost" showed that the summary was fine but possibly included some unimportant details. Four different colors were used to indicate different performance levels of a students' summary-writing in the four aspects. "Red" went with "poor"; "orange" with "fair"; "green" with both "almost" and "excellent". To view detailed feedback on the four aspects of summary-writing, the students could click on the four editing tools to look at the problematic sentences highlighted by the system. The grading criteria of the four aspects were shown below in Figure 3.

Copying

Status	Percentage of words that appear in the same order in the reading
Poor	20% or more
Fair	3 - 19%
Almost	0.1 - 2.9%
Excellent	0

Spelling

Status	Number of possible misspellings
Poor	8 or more
Fair	3 – 7
Almost	1 – 2
Excellent	0

Repeated

Status	Percentage of sentences that are highly similar in meaning
Poor	30% or more
Fair	5 – 29.9%
Almost	0.1 – 4.9%
Excellent	0

Unimportant

Status	Percentage of sentences that may be off-topic or not central to the reading
Poor	20% or more
Fair	10 – 19.9%
Almost	0.1 – 9.9%
Excellent	0

Figure 3. Grading criteria of "copying", "spelling", "repeated" and "unimportant".

Additionally, in the middle of the scoreboard, the researcher could give comments on students' summaries and make them viewable to the students. The students could click on the "Teacher Comments" button to read the comments and suggestions provided online by the teacher. By using the "note" function of the system, the researcher could provide comments and suggestions on any part of the students' summary, as such giving the students specific revision suggestions. By getting feedback from the researcher, the students would know that the researcher was keeping track of their performance, thus their motivation to compose a good summary could be further enhanced.

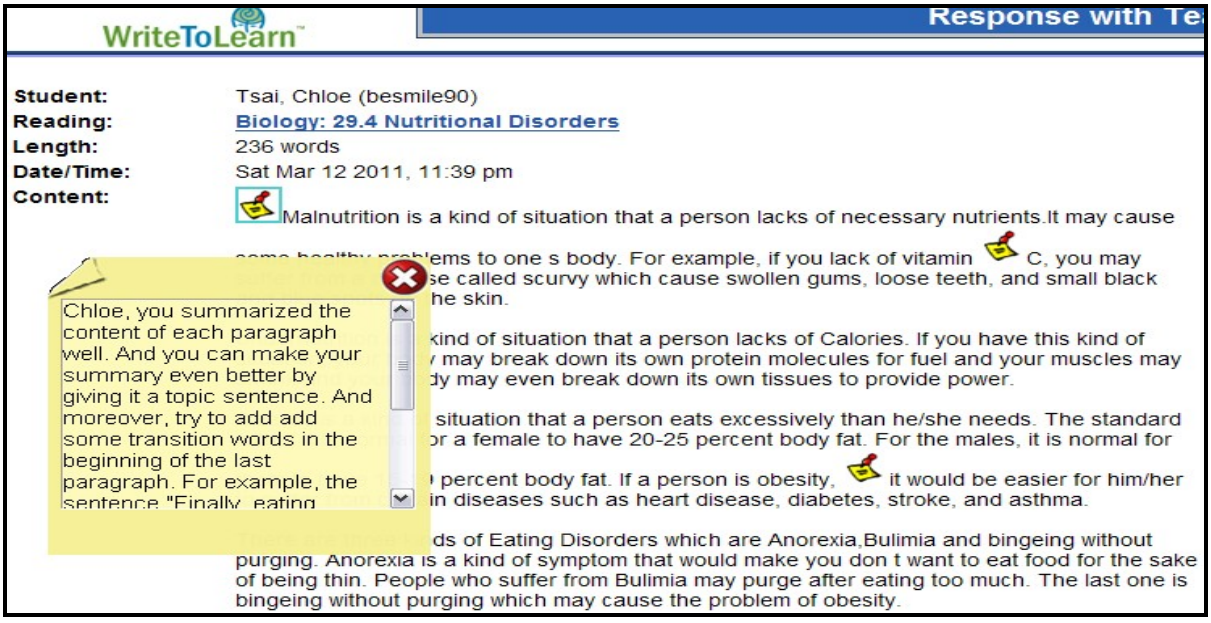

Figure 4. Sample of teacher comments.

The WriteToLearn training session lasted for 14 weeks (from week 1 to

week 14). During the training session, the participants were required to read a total of 10 passages (one passage a week excluding the mid-term exam week and the holidays) and write a summary for each passage. After submitting the first summary of each passage, the students were encouraged to revise their summaries based on the online feedback and teacher comments. For each passage, they were allowed to revise their summaries up to 5 times at most. The summary-writing assignments were counted toward the final grade of this course and weighed 10%. Students were told that summaries which reached the green zones in all different aspects would be awarded at least 80 points with 100 points as a maximum and only the best summary (out of the possible six drafts) was going to be counted.

The ten passages had been selected from WriteToLearn by the researcher according to their length and contents. The length of the articles selected ranges from 600 to 1000 words. The researcher chose the passages that might interest students. Half of the passages were of science-related topics and the other half were of social-science related topics. The titles of the passages were listed in Table 1. See Appendix 1 for a sample passage.

Table 1

Titles of Passages Selected for the Summary-training

Types of the Passage	Topics
Social-science-related Topics	• Freedom of Speech and Assembly in the United States • Ancient Greece and Modern Culture • People's Effect on the Environment • The United States – the Nation of Immigrants • The Atlantic Slave Trade
Science-related Topics	• Pandas • Nutrition Disorders Damage Health • Why Go To Mars? Getting to Know the Planet Nextdoor • Smoking Damages the Body and Shortens Life • Inside Earth: Earthquake Safety

3.3 Instruments

3.3.1 Reading Comprehension Test

To measure students' reading comprehension ability, the researcher used a multiple-choice reading comprehension test with two equivalent forms. Form A and Form B were administered respectively before and after the treatment. Form A and Form B had been adopted from both Tunghai on-line diagnostic reading test and the New TOEFL reading test. Each test contained four passages and 32 multiple-choice questions including 10 macro-level questions and 18 micro-level questions. The macro-level questions, as defined in the Review of Literature chapter, included main-idea identifying questions, between-paragraph inferencing questions, and summary-completing questions. The micro-level questions, as defined in the Review of Literature chapter, included detail comprehension questions, paragraph main-idea identifying questions, within-paragraph inferencing questions, questions asking for word meaning, and pronoun questions. Among the four passages, two were of science topics and the other two were of social-science topics. The topics of the passages were not directly related to the reading materials provided in the WriteToLearn program. See the sample test in Appendix 2. Table 2 and Table 3 list the example test items of the macro-level and micro-level questions.

Table 2

Example Test Items of Macro-level Questions

Types of macro-level questions	Example test items
Main idea identifying	What is the main idea of this passage?
Between-paragraph inferencing	What can be inferred from this passage about Great Britain before the Industrial Revolution?
Between-paragraph comprehension	All of the following are mentioned in paragraphs 1 and 2 as evidence of right-handedness in art and artists EXCEPT?
Summary-completing	An introductory sentence for a brief summary of the passage is provided below. Complete the summary by selecting the THREE answers that express the most important ideas in the passage. Some sentences do not belong in the summary because they express ideas that are not presented in the passage or are minor ideas in the passage.

Table 3

Example Test Items of Macro-level Question

Types of micro-level questions	Example test items
Detail comprehension	According to the passage in paragraph 2, why must a person touch a torpedo in two places to get a shock?
Paragraph main-idea identifying	What is the main idea of paragraph 5?
Within-paragraph inferencing	Which of the following statements about fractures and cut marks can be inferred from paragraph 4?
Word-meaning identifying	The phrase **depicted** in paragraph 2 is closest in meaning to?
Pronoun reference	What does **they** refer to in paragraph 3?

Form A and Form B both contained 28 multiple-choice questions. Twenty-six of the questions were worth 1 point each and contain 4 options in each question. The other 2 questions count for 3 points each, for the two questions required more complex reading-comprehension skills, like completing a summary by selecting the main points of an article from several choices. Thus the full score of RC Test (either Form A or Form B) is 32(1 points $\times$26 + 3 points $\times$2).

To counterbalance the possible different difficulty level of Form A and Form B, in the pretest, the researcher had half of students in the high-level and low-level reading groups take Form A and the other half Form B. In the posttest, those who had taken Form A in the pretest took Form B and those who had taken Form B in the pretest took Form A. The pretest was administered in week 3, before the WriteToLearn training session, and the posttest was administered in week 15, after the WriteToLearn training was finished.

3.3.2 Vocabulary Knowledge Test

To measure students' vocabulary knowledge growth, the researcher developed a vocabulary knowledge test. The format of the test was adopted from the Vocabulary Knowledge Scale (VKS) (Paribakht & Wesche, 1993). The Vocabulary Knowledge Scale was initially developed for ESL learners in a university setting, and research has shown that VKS was able to capture changes in vocabulary-knowledge growth during a short period of instructional experiment. For each word tested in Vocabulary Knowledge Test, the students had to choose from the following scale, as shown in Table 4, according to their knowledge of the word.

Table 4

Vocabulary Knowledge Scale Categories

Number of Category	Description of knowledge of the target word
I.	I don't remember having seen this word before.
II.	I have seen this word before, and I think it means _____. (synonym or translation)
III.	I have seen this word before, and I think it means _____. (synonym or translation)
IV.	I know this word. It means _____. (synonym or translation)
V.	I can use this word in a sentence: _____. (If you do this section, please also do Section 4.)

The maximum point of VKT is 5 and the minimum point is 1. Table 5 below summarizes the scoring criteria. For each word, when a student chose Category I or II, one or two points were given. If Category III was chosen and there was a correct synonym or translation, the student could obtain 3 points. If Category III was chosen but the synonym or translation was wrong, only two points were given. Wrong answers in Category IV and V also led to only 2 points. If in Category V a student gave a correct meaning of the word but was inappropriate for the context, 3 points were given. If the student gave a correct meaning of the word for the sentence context but had grammatical errors, 4 points were given. In Category V, a student would get 5 points if he or she gave a sentence which showed both correct semantic meaning for the context and correct grammatical usage.

Table 5

Scoring Criteria of Vocabulary Knowledge Test

The item chosen	Types of answers	Scores obtained
I. I don't remember having seen this word before.		1
II. I have seen this word before, but I don't know what it means.		2
III. I have seen this word before, and I think it means _____. (synonym or translation)	Correct synonym or translation provided	3
	Incorrect translation or synonym provided	2
IV. I know this word. It means _____. (synonym or translation)	Correct translation provided	4
	Incorrect translation or synonym provided	2
V. I can use this word in a sentence: _____.	Correct translation provided, but the translation is inappropriate for the context	3
	Correct meaning of the word for the sentence context provided, containing grammatical errors	4
	Correct semantic meaning of the word for the sentence context provided, along with correct grammatical usage.	5

For the present study, a total of 52 words were tested in the Vocabulary Knowledge Test. They were chosen from the ten WriteToLearn passages for which the participants were required to write summaries during the WriteToLearn training session. In the test development stage, the researcher had selected several words from each passage which would possibly be unknown and hence would pose some difficulties to the participants. Words that were possibly unknown to the participants were usually low-frequency words. Since the participants were freshmen in college, the researcher supposed that they might have learned most of the words in the GEPT Intermediate Level Word List before being admitted to college. The GEPT

Intermediate Level Word List includes words at and below the 4000 word level as compiled by the Language and Training Test Center. Therefore, when selecting words for the Vocabulary Knowledge Test, the researcher tended to look at words which are beyond the 4000 word level. In addition, the researcher also paid attention to words which are important content words (verbs, nouns or adjectives). Words that have semantic relationship with multiple words in the surrounding were most likely to be important words and the participants should have greater opportunities to pick up the meanings of the words through reading. For instance, the word "reinforce" was selected from the following sentence in a passage about earthquake.

> *During an earthquake, brick buildings and some wood-frame buildings may collapse if their walls have not been **reinforce**d, or **strengthen**ed.*

The word "reinforce" was included in the Vocabulary Knowledge Test because it is a lower frequency word not listed in the GEPT Intermediate Level Word List. Moreover, it is semantically related to "collapse" and 'strengthen" and hence is an important content word.

Through this procedure, a total of eighty words were chosen for the pilot test. The vocabulary knowledge test was piloted with 12 sophomore English-majors. Their performances were analyzed. The research followed the scoring procedure mentioned above and calculated the average score. Words that had an average score higher than 3 were excluded from the test because an average of 3 points implied most learners had known the word. In addition, the researcher also excluded words that have multiple meanings because those words had only one meaning in the passages where they appeared. Finally, 28 words were excluded and 52 words were kept for the Vocabulary Knowledge Test. See Appendix 3 for a complete list of the 52 words. The full score of the test was hence 260 points (5 points x 52 words) and the minimum score was 52 points (1 points x 52 words). The Vocabulary Knowledge pretest was administered in week 3, before the WriteToLearn training session, and the posttest was administered in week 15, after the WriteToLearn training was completed.

3.4 Data Collection

The data collection procedure is summarized in Table 6. The WriteToLearn training session lasted for 14 weeks. In the first week of the spring semester of 2011, the participants first received summary-writing instructions, practiced writing a summary in class, and discussed on their summaries with peers. In week 3, the participants took the reading-comprehension pretest and the Vocabulary Knowledge pretest, and were introduced to the WriteToLearn online system. From week 4 to week 14, all the participants were required to do an online summary-writing assignment each week on WriteToLearn, and completed 10 summaries in total (the midterm week being excluded). During the training, students' summary-writing performances were recorded by the program. Finally, the reading -comprehension posttest and Vocabulary Knowledge posttest were administered in week 15.

Table 6

Date Collection Procedure of WriteToLearn Summary-writing Training

Time	Instructions, Treatments, and Test Administration
Week 1	Summary-writing instructions
Week 3	Reading comprehension pretest Vocabulary Knowledge pretest
Week 4 ~ Week 14	WriteToLearn online summary writing
Week 15	Reading comprehension posttest Vocabulary Knowledge posttest

3.5 Data Analysis

To see whether the two groups of EFL learners (high-level and low-level) repectively improved their reading comprehension and vocabulary knowledge upon the completion of the *WriteToLearn* program, the researcher conducted dependent t-tests to see if each group respectively made significant improvements in reading comprehension and vocabulary knowledge after finishing the summary-writing training. The researcher also conducted

independent t-tests to see if there were any significant differences between the posttest performances of the two groups on reading comprehension as well as vocabulary knowledge.

Additionally, in order to depict the summary-writing performances of the two groups of EFL learners in the ten summary-writing assignments, the learners' profiles of summary writing assignments as recorded in the WriteToLearn program were examined. Percentages of the four writing aspects (copying behavior, spelling mistakes, repeated content and unimportant content) in the participants' first summary drafts were calculated. Comparisons were made between the high-level and the low-level reading groups.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Results of the Data Analyses

The present study implemented an on-line learning program called WriteToLearn in a reading class. The program was designed to facilitate reading comprehension development through summarization of reading passages. Thus, the participants were required to read the on-line articles, compose summaries and submit them to receive automated feedback. The researcher examined the effects of the on-line program by looking at the participants' improvements in reading comprehension, vocabulary knowledge, and summary writing. The following sections describe the participants and their improvements in reading comprehension, vocabulary knowledge and summary writing over the 14-week training session.

Since in the present study the researcher aimed at examining to what extent the participants at two ability levels benefited from the on-line program, the participants were divided into two groups -- high-level reading group (n=17) and low-level reading group (n=16), mainly based on their final exam scores of the reading course in the previous semester (i.e. fall semester). Minor adjustments were made based on their pretest reading-comprehension scores. Table 7 shows the distributions of the reading-comprehension pretest scores in the high-level and low-level reading groups.

Table 7

Distribution of Reading-Comprehension Pretest Scores (max.=32 points) in Both Groups

Score	＜10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18
Number of participants									
High-level (n=17)	0	0	0	0	0	0	0	3	7
Low-level (n=16)	0	1	2	1	3	2	5	1	1

Score	18-19	19-20	20-21	21-22	22-23	23-24	24-25	＞25
Number of participants								
High-level (n=17)	2	0	2	1	1	0	1	0
Low-level (n=16)	0	0	0	0	0	0	0	0

4.1.1 Reading Comprehension (RC)

The study employed a pretest-posttest design when measuring the reading-comprehension improvement of the participants. In order to reduce the possible practice effect of using the same form for the posttest, the researcher had the participants take different 2 forms with one form on the pretest and the other on the posttest. Based on their final exam scores of the reading course from the previous semester, the researcher rank- ordered them from the highest to the lowest and numbered them. Those whose scores were above the median tentatively belonged to the high-level reading group; whereas those whose scores were below the median belonged to the low-level reading group. Then, those who were assigned an odd number took Form A in the pretest and Form B in the posttest; whereas those who were assigned even number took Form B in the pretest and Form A in the posttest.

The pretest score distribution was similar to the previous semester's final exam score distribution. After all the participants received the reading-comprehension pretest, the researcher adjusted the grouping according to the pretest scores (Table 7).

The following paragraphs describe the test results and the comparisons of the performances of the two groups on the reading comprehension test. All hypothesis testing was conducted at the significance level of .05 (i.e. $\alpha = .05$).

Table 8 shows the means and standard deviations of the RC pretest and posttest scores of the high-level and low-level reading groups. The mean score of the high-level reading group on the RC pretest was 18.45, with a standard deviation of 2.42. The mean score of this group on the posttest was 18.67, with a standard deviation of 3.59. As for the low-level reading group, their mean score on the RC pretest was 13.84, with a standard deviation of 1.99. Their mean score of the posttest was 16.94, with a standard deviation of 2.17. As can be seen here, the high-level reading group outperformed the low-level reading group on the pretest, but scored just slightly higher than the low-level reading group on the posttest. To determine if a statistically significant difference existed between the two groups in the pretest and posttest scores, two independent t-tests were conducted. As shown in Table 9 and Table 10, there was significant difference between the two groups' pretest scores (t= 5.944, p= .000). The high-level reading group scored significantly higher than the low-level reading group in the pretest. However, in the posttest, no significant difference was found between the two groups' scores (t= 1.665, p= .106).

Table 8

Means and Standard Deviations of the Reading-Comprehension Pretest and Posttest Scores of the High-level Reading Group and Low-level Reading Group

Group	n	Mean	SD
High-level			
Pretest	17	18.45	2.42
Posttest	17	18.67	3.59
Low-level			
Pretest	16	13.84	1.99
Posttest	16	16.94	2.17

Note. Possible maximum score is 32

Table 9

The Mean Differences of the Two Groups' Reading-Comprehension Pretest

	Mean difference (High level scores– Low level scores)	Standard error	t-value	df	p-value
Pretest	4.61	0.775	5.944	31	.000

Table 10

The Mean Differences of the Two Groups' Reading-Comprehension Posttest

	Mean difference (High level scores– Low level scores)	Standard error	t-value	df	p-value
Poesttest	1.73	1.041	1.665	31	.106

Additionally, as shown in Table 8, the reading comprehension of the high-level reading group did not seem to improve upon the completion of the WriteToLearn training. In contrast, the reading comprehension of the low-level reading group seemed to have improved to a greater extent. To determine if a statistically significant difference existed between the pretest and posttest scores within each group, two dependent t-tests were conducted for each group. As can be seen in Table 11, for the high-level reading group, the results show that there is no significant difference between the pretest and posttest scores (t= 0.192, p= .850). That is, the high-level reading group made no significant improvement in their reading-comprehension performance through the WriteToLearn training session. For the low-level reading group, the results show that there is significant difference between the pretest and posttest score (t= 4.420, p=.000). That is, the low-level reading group has significantly improved their reading-comprehension after the WriteToLearn training session. In sum, the high-level reading group did not improve in their reading-comprehension throughout the WriteToLearn training session, while the low-level reading group made a significant improvement after the

WriteToLearn training.

Table 11

The Mean Differences of the Two Groups' Reading-Comprehension Pretest and Posttest

	Mean difference (Posttest scores -Pretest scores)	Standard error	t-value	df	p-value
High-level	0.224	1.162	0.192	16	.850
Low-level	3.100	0.701	4.420*	15	.000

The Reading Comprehension Test in this study consisted of two kinds of questions: micro-level reading questions and macro-level reading questions. To further examine whether the participants made improvement in micro-level reading comprehension and macro-level reading comprehension, apart from analyzing the two groups' overall scores on RC Test, the researcher further examined the two groups' performance in micro-level questions and macro-level questions separately.

Table 12 shows the two groups' mean scores and standard deviations of micro-level questions and macro-level questions in both pretest and posttest. For the high-level reading group, the pretest mean score of micro-level questions is 11.00, with a standard deviation of 1.80. The pretest mean score of micro-level questions is 11.12, with a standard deviation of 2.37. As for the macro-level questions, the pretest mean score in pretest is 7.45, with a standard deviation of 1.52, while the posttest mean score is 7.55, with a standard deviation of 1.66. As shown in the statistical results, compared to the pretest scores, the high-level reading group did not seem to have any obvious improvement in both micro-level scores and macro-level scores on the posttest.

As for the low-level reading group, their pretest mean score of micro-level questions is 7.94, with a standard deviation of 1.34. In the posttest, their mean score in micro-level questions is 10.06, with a standard deviation of 1.57. As for the low-level reading group's performance on macro-level questions, the pretest mean score is 5.90, with a standard deviation

of 1.36. The posttest mean score of macro-level questions is 6.88, with a standard deviation of 1.65. The low-level reading group seemed to have performed much better on both micro-level and macro-level questions in the posttest than in the pretest.

Table 12

Means and Standard Deviations of the Micro-level and Macro-level Scores of Both Groups in Reading-comprehension Pretest and Posttest

	Group	Micro-level questions [a]		Macro-level questions [b]	
		Mean	SD	Mean	SD
Pretest	High	11.00	1.80	7.45	1.52
	Low	7.94	1.34	5.90	1.36
Posttest	High	11.12	2.37	7.55	1.66
	Low	10.06	1.57	6.88	1.65

Note. [a]Maximum score is 18. [b]Maximum score is 14.

The researcher conducted two dependent t-tests to see whether statistically the low-level reading group made significant improvement in micro-level and macro-level questions in the RC posttest. As can be seen in Table 13, the result indeed showed that there is a significant difference between the low-level reading group's micro-level pretest scores and posttest scores (t=4.490, p=.000). There is also a significant difference between the low-level reading group's macro-level pretest scores and posttest scores (t= 2.203, p=.044). That is, the low-level reading group improved significantly on both micro-level and macro-level questions throughout the WriteToLearn training session.

Similarly, the researcher also conducted dependent t-tests to see if there is any significant difference between the pretest and posttest performance of the high-level reading group. As shown Table 13, the results show that there is no significant difference between the high-level reading group's pretest and posttest micro-level scores (t= 0.135, p=.895). Also, no significant difference is found between the pretest and posttest macro-level scores (t= 0.188, p=.853). That is, the high-level reading group had improved neither on macro-level questions nor micro-level

questions after the WriteToLearn training.

Table 13

The Mean Differences of the Two Groups' Micro-level and Macro-level Score in Reading-Comprehension Pretest and Posttest

	Mean difference (Postest scores – Pretest scores)	Standard error	t-value	df	p-value
High-level					
Micro	0.118	0.874	0.135	16	.895
Macro	0.106	0.563	0.188	16	.853
Low-level					
Micro	2.125	0.473	4.490*	15	.000
Macro	0.975	0.443	2.203*	15	.044

4.1.2 Vocabulary Knowledge

To investigate whether the two groups of participants improved in their vocabulary knowledge through the WriteToLearn training session, the researcher had both of the two groups take the Vocabulary Knowledge Test as pretest before the WriteToLearn training and as posttest after the WriteToLearn training. Fifty-two words were selected from the ten articles which were assigned for the learners to read and compose summaries during the training. For each word, the maximum score is 5 points and the minimum is 1. The full score of the Vocabulary Knowledge Test is 260 points (5 points × 52 words) and the minimum score is 52 points (1 point × 52 words).

Table 14 shows the pretest and posttest scores of the two groups in Vocabulary Knowledge Test. The pretest mean score of the high-level reading group is 121.65, with a standard deviation of 26.67.The pretest mean score of the low-level reading group is 109.75, with a standard deviation of 19.92. As shown in Table 15, the independent T-test result shows that there is no significant difference between the two groups' Vocabulary Knowledge pretest scores (t= 1.445, p=.159).

Table 14

Means and Standard Deviations of the Two Group's Vocabulary Knowledge Scores on Pretest and Posttest

Group	n	Mean	SD
High-level			
Pretest	17	121.65	26.67
Posttest	17	146.86	31.74
Low-level			
Pretest	16	109.75	19.92
Posttest	16	132.69	26.10

Note. Maximum score = 260

Table 15

The Mean Differences of the Two Groups' Vocabulary Knowledge Pretest Scores

	Mean difference (High level scores– Low level scores)	Standard error	t-value	df	p-value
Pretest	11.897	8.235	1.445	31	.159

As for whether the two groups made improvements in their vocabulary knowledge after the WriteToLearn training, their performances are discussed as follows.

The high-level reading group had a mean score of 121.65 on Vocabulary Knowledge pretest, while their mean score on posttest was 146.86. The difference between the pretest and posttest mean scores shows that the high-level reading group seems to have made some improvement in their vocabulary knowledge on the posttest. As can be seen in Table 16, the dependent T-test further indicates that the difference between the pretest and posttest scores was significant (t= 9.170, p=.000). In other words, upon the completion of the WriteToLearn training, the high-level reading group made a statistically significant improvement in their vocabulary knowledge performance.

The low-level reading group also had made a significant improvement in their vocabulary knowledge. Their pretest mean score was 109.75 and their posttest mean score was 132.69. As can been seen in Table 16, the dependent t-test result shows that the difference between the

pretest and posttest scores of the low-level reading group was significant (t= 6.899, p=.000). That is, similar to the high-level reading group, the low-level reading group also improved significantly on the Vocabulary Knowledge Test upon the completion of the WriteToLearn training.

Table 16

The Mean Differences of Each Group's Vocabulary Knowledge Pretest and Posttest

	Mean difference (Posttest scores – Pretest scores)	Standard error	t-value	df	p-value
High-level	25.177	2.746	9.170*	16	.000
Low-level	22.936	3.325	6.899*	15	.000

The t-test results mentioned above show that there is no significant difference between the high-level and low-level reading groups' performances on Vocabulary Knowledge pretest. Upon the completion of the WriteToLearn training, both the high-level and low-level reading groups have made significant improvements on Vocabulary Knowledge posttest. Furthermore, to see whether the WriteToLearn training had different effects on the vocabulary knowledge performance of the two groups, the researcher conducted independent t-test on the two groups' Vocabulary Knowledge posttest scores. As can be seen in Table 17, no significant difference was found (t= 1.392, p=.174). That is, there is no significant difference between the two groups' Vocabulary Knowledge posttest scores.

Table 17

The Mean Differences of the Two Groups' Vocabulary Knowledge Posttest Scores

	Mean difference	Standard error	t-value	df	p-value
Posttest	14.136	10.153	1.392	31	.174

Although the high-level reading group scored higher than the low-level reading group in both Vocabulary Knowledge pretest and posttest, the statistical analysis results show that there is no significant difference between the performances of the two groups in Vocabulary

Knowledge pretest, and nor is there significant difference in the posttest. However, both of the groups have made a statistically significant improvement in Vocabulary Knowledge posttest as compared to their pretest performances.

4.1.3 Summary-Writing

To examine whether the two groups of participants (high-level and low-level) respectively made progresses in their summary writing during the 14-week training session using WriteToLearn and if there was any significant differences between the two groups in their progresses of summary writing, the researcher did an analysis on the feedback generated by WriteToLearn system on the participants' summaries. The feedback includes four aspects: copying behavior, spelling mistakes, repeated content and unimportant content, all of which should be avoided in any good summaries.

"Copying" counts the percentage of words that are taken directly from the original passage instead of being paraphrased or rewritten by the writer's own words. "Spelling" shows the spelling mistakes in a summary. "Repeated" gives an estimate of sentences that are similar in meaning, pointing out the unnecessary repetitive content in a summary. "Unimportant" reflects the percentage of irrelevant details that appear in a summary. Each time a participant submitted a summary onto the WriteToLearn system, the system provided the four aspects of feedback in the form of remarks (Excellent, Almost, Fair or Poor) along with specific percentages. The lower the percentages are, the better the quality of the summary is, and vise versa. When analyzing the feedback on the participants' summary-writing, the researcher focused on their first drafts of all ten summaries. For each aspect of the feedback, the average percentage of each group was calculated and presented in tables as well as in charts. The following paragraphs describe the participants' overall performance first and then discuss each aspect in more details.

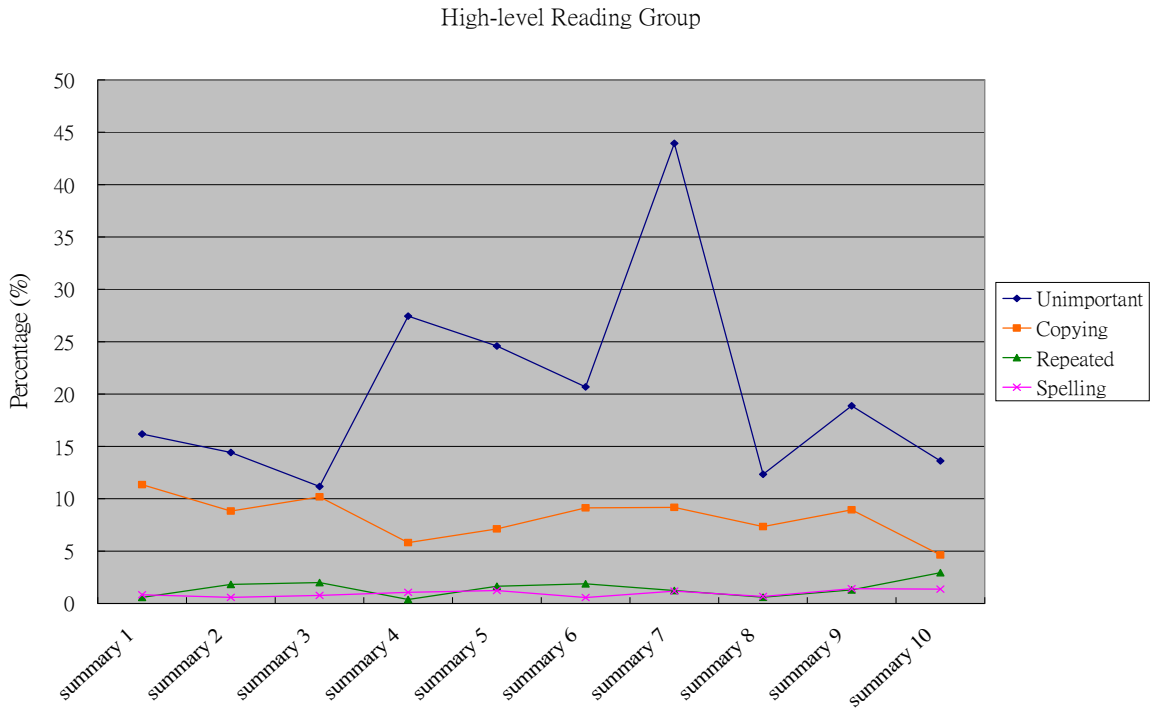

Figure 5. Percentage of copying, spelling, repeated and unimportant in the high-level reading group throughout the 10 summary-writing assignments

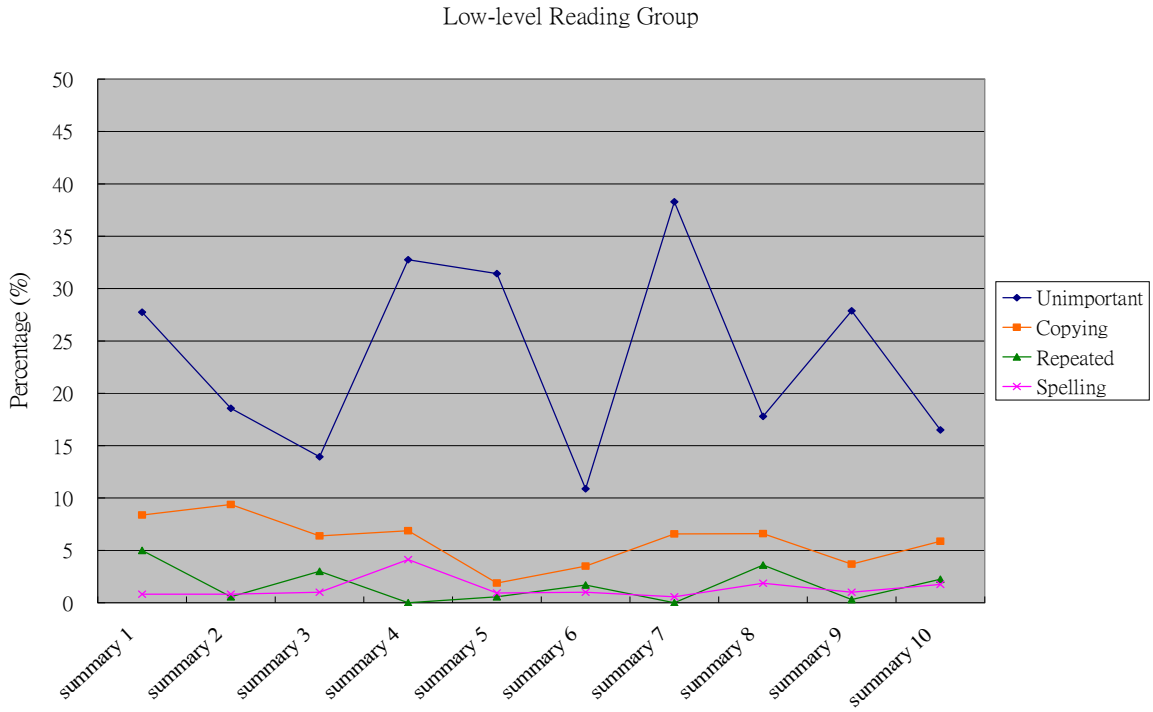

Figure 6. Percentage of copying, spelling, repeated and unimportant in the low-level reading group throughout the 10 summary-writing assignments

Figure 5 and Figure 6 show the group average percentage of copying behavior, spelling mistakes, repeated content, and unimportant content of the summary drafts calculated by the

WriteToLearn system. Figure 5 shows the high-level reading group's performance and Figure 6 shows the low-level reading group's. The lower the average percentage was, the better the average performance of the group was. First of all, as can be seen in the two figures, both groups had quite low percentage in spelling mistakes and repeated content, always no higher than 5 percent. And the lines for spelling mistakes and repeated content of both high-level and low-level reading groups swang between 0 to 5 through the 10 summaries.

Second, the lines for copying behavior of both high-level and low-level reading groups showed a slight decline throughout the ten summaries. The percentages of copying behavior seemed to gradually decline throughout the summary-writing training session. Lastly, the lines of unimportant content went up and down, which implied that the learners' performance was probably greatly impacted by the differences among the passages.

In the next section, the two groups' performances in copying behavior, spelling mistakes, repeated content and unimportant content are compared to each other respectively.

Table 18 shows the group mean percentages of copying behavior throughout the ten summaries. As can be seen in Figure 7, both groups showed a tendency of declining in the copying behavior throughout the ten summary-writing activities. Interestingly, the low-level reading group generally demonstrated lower percentage of copying behavior than the high-level reading group. In most summaries, the low-level reading group seemed to be more capable of using their own words when composing their summaries and hence performed better than the high-level reading group in the "Copying" aspect.

Table 18

Group Mean Percentage of Copying Behavior in Each Summary Assignment

Group	1	2	3	4	5	6	7	8	9	10
High-level	11.35%	8.82%	10.18%	5.81%	7.12%	9.13%	9.18%	7.33%	8.94%	4.63%
Low-level	8.38%	9.38%	6.38%	6.88%	1.88%	3.50%	6.57%	6.60%	3.69%	5.88%

\

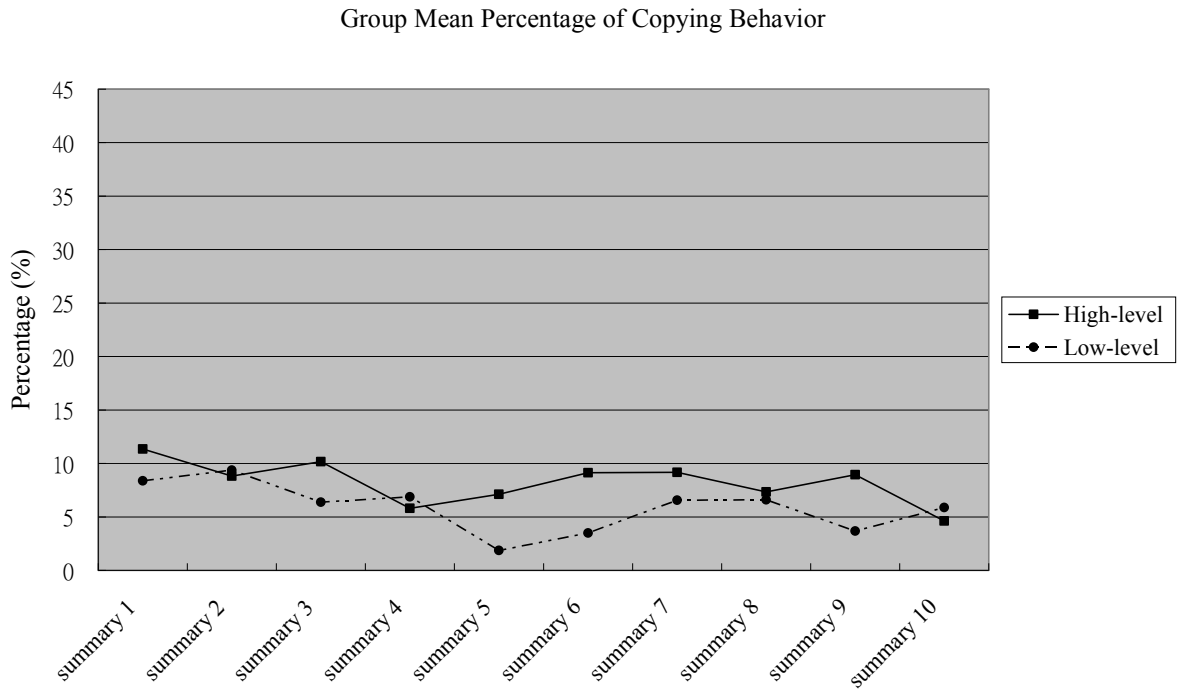

Figure 7. Group mean percentage of copying behavior

Table 19 shows the group mean percentage of spelling mistakes in each summary. Figure 8 shows that except for summary 4 by the low-level reading group, the percentages of spelling tended to slightly increase through the ten summary-writings. In other words, the participants did not improve in word-spelling throughout the WriteToLearn training session, but made slightly more spelling mistakes. However, except for the fourth summary, the "Spelling" percentages of both groups stayed below 2 percent, which means that the two groups of participants only made few spelling mistakes from the first to the tenth summary. Therefore, despite the slightly increasing percentage of spelling mistakes during the training session, both the high-level and low-level reading groups made few spelling mistakes throughout the ten

summaries.

Table 19

Group Mean Percentage of Spelling Mistakes in Each Summary Assignment

	1	2	3	4	5	6	7	8	9	10
High-level	0.8%	0.6%	0.8%	1.1%	1.2%	0.6%	1.2%	0.7%	1.4%	1.4%
Low-level	0.8%	0.8%	1.0%	4.1%	0.9%	1.0%	0.6%	1.9%	1.0%	1.8%

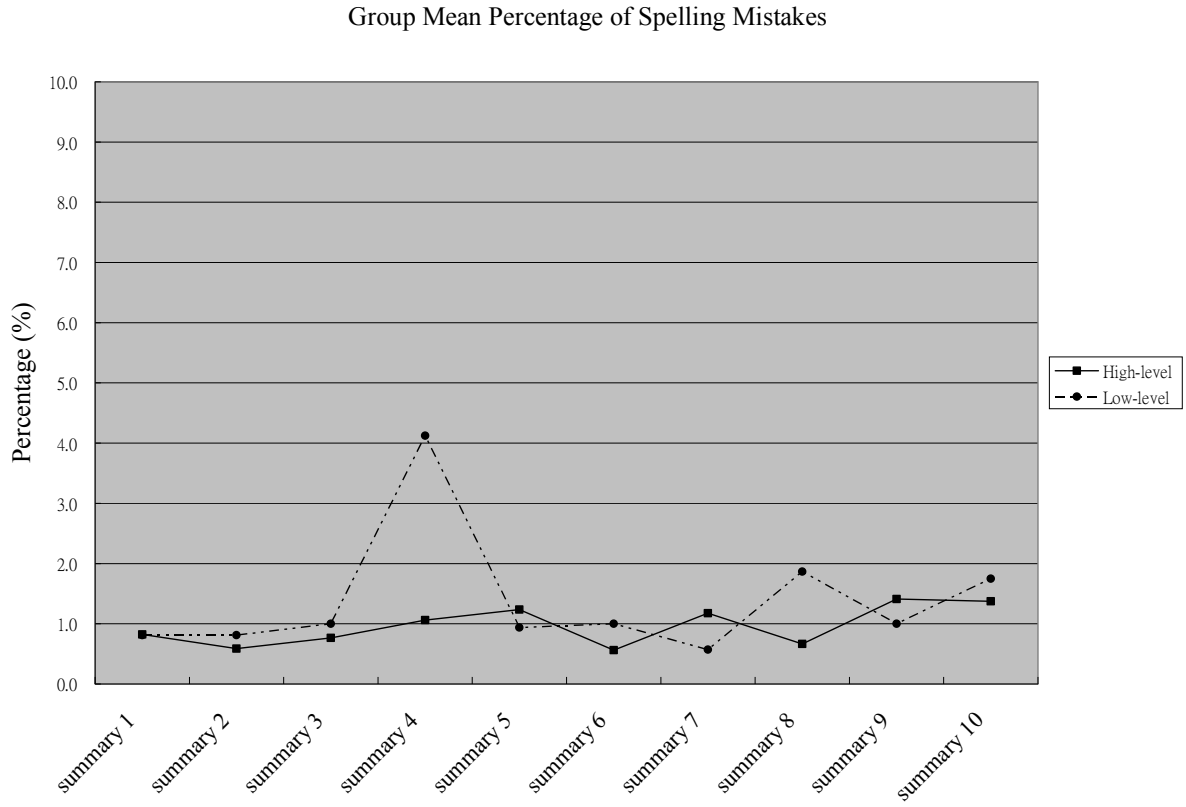

Figure 8. Group mean percentage of spelling mistakes

Table 20 shows the group mean percentages of repeated content of the two groups from the first to the tenth summary. As shown in figure 9, the two lines went up and down constantly from the first summary to the tenth. However, similar to the lines of spelling mistakes discussed above, the two groups' lines of repeated content moved up and down slightly between 0 to 2 percent, which can be considered very low percentage. That is, despite the ups and downs of the two "Repeated" lines, both the high-level and low-level reading groups generally had very low percentages (0-5%) of repeated content in the ten summaries.

Table 20

Group Mean Percentage of Repeated Content in Each Summary Assignment

	1	2	3	4	5	6	7	8	9	10
High-level	0.59%	1.82%	2.00%	0.38%	1.65%	1.88%	1.24%	0.60%	1.29%	2.94%
Low-level	5.00%	0.56%	3.00%	0.00%	0.56%	1.69%	0.00%	3.60%	0.31%	2.25%

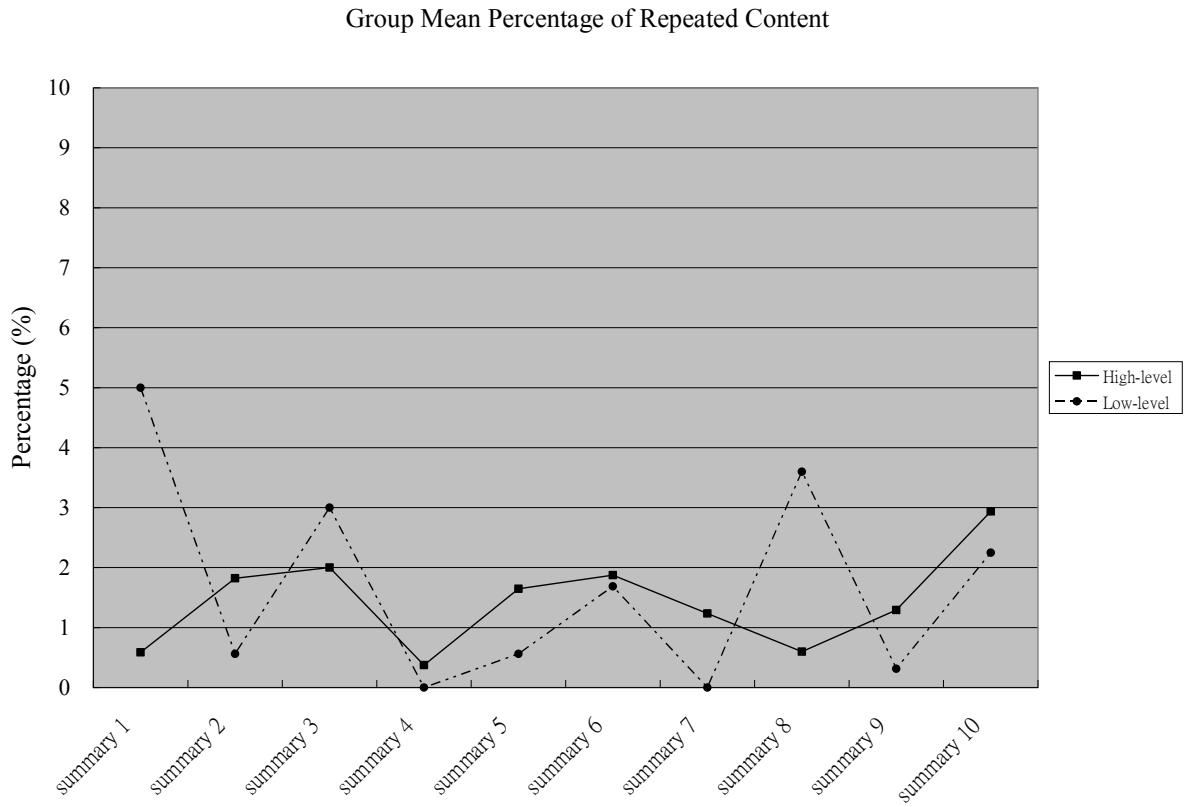

Figure 9. Group mean percentage of repeated content

Table 21 shows the average "Unimportant" percentages of the two groups from the first to the tenth summary. The unimportant content refers to the amount of irrelevant content or details that existed in the participants' summaries. Figure 10 indicates that there seems to be no obvious improvement in the two groups' performances in the "Unimportant" aspect. The shapes of the two groups' lines were very similar. The two groups had lower percentages of "Unimportant" content in summary number 3, 6, 8 and 10, while they both had higher percentages in summary number 4, 7 and 9. The movement of the lines again might be due to the differences in text lengths and possible content difficulties of the passages. The results

imply that the participants' performances in grabbing important content and leaving out the irrelevant details were not steady.

Table 21

Group Mean Percentage of Unimportant Content in Each Summary Assignment

	1	2	3	4	5	6	7	8	9	10
High-level	16.18	14.41	11.18	27.44	24.59	20.69	43.94	12.33	18.88	13.63
Low-level	27.75	18.56	13.94	32.75	31.44	10.88	38.29	17.80	27.88	16.50

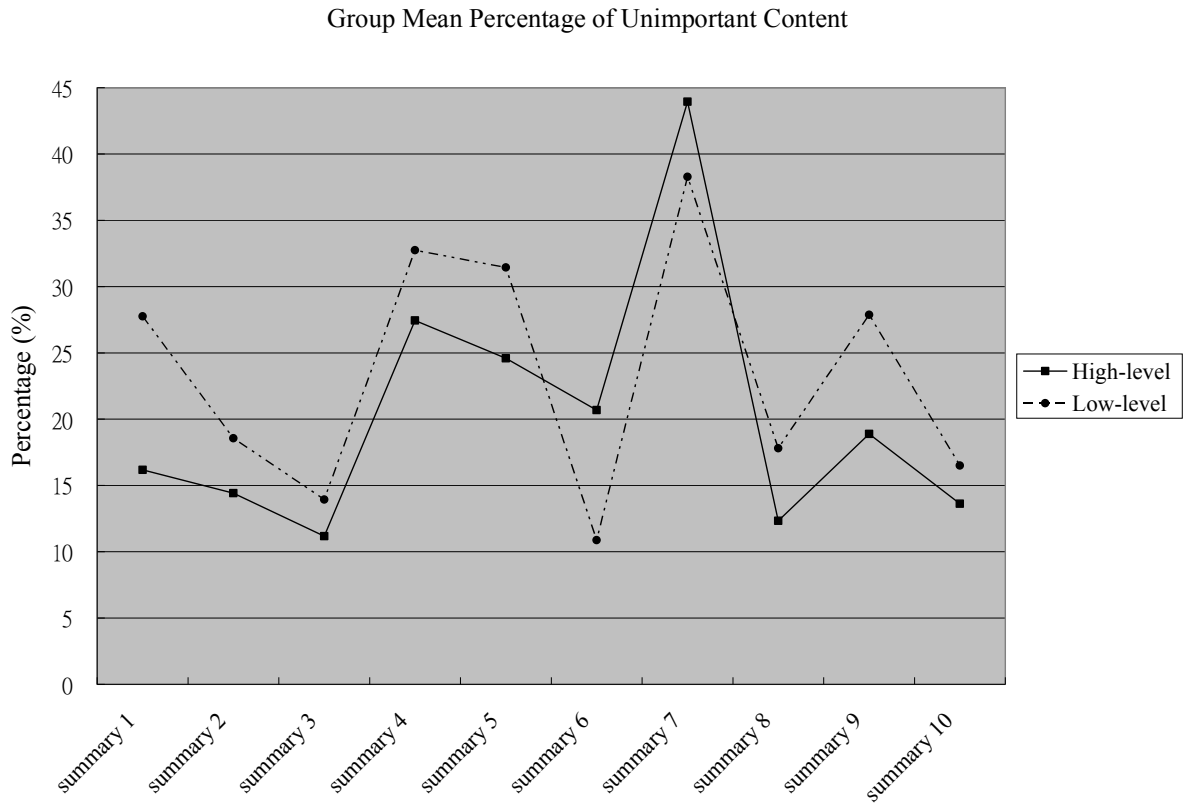

Figure 10. Group mean percentage of unimportant content

It has been demonstrated in this chapter the performance of the high-level and low-level reading groups before and after the WriteToLearn training session in reading comprehension, vocabulary knowledge and summary writing. To sum up, as for the reading comprehension performance, the high-level group outperformed the low-level group on the pretest but not on the posttest. In addition, the high-level group did not show any improvement in reading comprehension upon the completion of the summary-writing training, while the low-level

group made a significant progress. As for the vocabulary knowledge performance, there was no significant difference between the high-level and low-level groups' performances on the Vocabulary Knowledge pretest, nor was there significant difference on the posttest. Both group made significant improvements in Vocabulary Knowledge Test upon the completion of the summary-writing training.

Finally, in terms of their summary-writing performances, upon the completion of WriteToLearn training, both high-level and low-level reading groups showed slight progresses in copying behavior. Both groups had very low percentages of spelling mistakes and repeated content throughout the summary-writing training. As for the unimportant content, the two groups did not seem to make obvious progress throughout the WriteToLearn training. Overall, throughout the training both groups of participants seemed to improve in the "Copying" aspect and seemed not to make progresses in the "Unimportant" aspect, and their performance in "Spelling" and "Repeated" aspects remained steadily good throughout the training session.

4.2 Discussion

4.2.1 Reading-Comprehension (RC)

The reading-comprehension test results show that the high-level reading group did not improve in their RC performance upon the completion of the WriteToLearn training, while the low-level reading group made a significant improvement after the WriteToLearn training. That is, in terms of facilitating learners' development of reading comprehension, the WriteToLearn online summary-writing system seemed to have benefited the low-level reading group more than the high-level reading group.

To explain the little improvement of the high-level reading group in their RC performance, the researcher examined the feedback of the WriteToLearn system and found that more high-level reading participants passed the default threshold of covering the important content in each paragraph than the low-level reading participants in most summary assignments. The

program's feedback on the section coverage of the participants' summaries is shown in Table 22. The WriteToLearn program separated the original article into several sections. The feedback the WriteToLearn system generated reflected the extent to which a summary covers the important content in each section. The feedback of section coverage can be categorized into 3 types: passing above the threshold, moderately falling below the threshold, and falling far below the threshold. Table 22 shows the number of participants who received "passing above the threshold" in all sections in their first drafts of each summary assignment. As can been seen, except for summary 6 and summary 7, more high-level reading participants passed the threshold in all sections than low-level reading participants in all the other summaries. Possibly because more of the high-level reading participants had met the basic requirements of the system in their first drafts of the summary assignments in comparison with the low-level reading participants, they did not need to try as hard as the low-level reading participants to reread the original articles and grab the important content throughout the summary-writing training process. Thus, compared to the low-level reading group, who had to try hard to find the important content for the purpose of meeting the requirement of section coverage, the high-level reading group did not benefit as much from WriteToLearn in terms of their reading comprehension ability.

Table 22

Number of Participants Passing the Threshold of Section Coverage in All the Sections (First draft)

Number of summary	1	2	3	4	5	6	7	8	9	10	Average
High-level reading group	1	4	11	5	11	6	7	6	1	5	5.7
Low-level reading group	0	0	6	2	4	7	7	2	0	3	3.1

Previous studies claim that summary-writing training facilitates learners' reading comprehension (Friend, 2001; Oded & Walters, 2001; Radmacher & Latosi-Sawin, 1995). The

present study further found that the WriteToLearn summary-writing training benefited EFL college learners at a relatively low level in terms of facilitating their reading comprehension performance. The summary-writing process required learners to put emphasis on knowledge of macrostructure, which is more crucial for text comprehension than knowledge of microstructure, and thus effectively benefited their reading comprehension. Moreover, the low-level reading participants in the present study had improved significantly on macro-level as well as micro-level comprehension upon the completion of WriteToLearn training. In other words, the WriteToLearn summary-writing training facilitated not only the EFL learners' macro-level comprehension but also their micro-level comprehension.

4.2.2 Vocabulary Knowledge

To examine whether the high-level and low-level reading group respectively improved in their vocabulary knowledge upon the completion of the WriteToLearn summary-writing training, the researcher had the two groups of participants do the Vocabulary Knowledge Test one time prior to the WriteToLearn training session and the second time upon the completion of the WriteToLearn training. No significant difference was found between the two groups' pretest scores nor between the two groups' posttest scores. However, what is worth mentioning is the significant improvements of the two groups in their Vocabulary Knowledge posttest performance compared to their pretest scores.

The 52 target words tested on the Vocabulary Knowledge Test had been selected from the ten articles that the participants summarized during the WriteToLearn training session. Moreover, based on the vocabulary-knowledge pilot test, the words possibly known by the participants before the training session were excluded. Therefore, reasonably no significant difference was found in the two groups' pretest scores. Upon the completion of the WriteToLearn summary-writing training, the high-level reading group as well as the low-level reading group made a significant improvement in vocabulary knowledge. While the result implies that the participants learned the words through the summary-writing process, the

limitation of the measurement may have led to the result. The Vocabulary Knowledge Test in this study was designed to measure the participants' depth of vocabulary knowledge, hence it was sensitive to the participants' familiarity of the target words. For each of the 52 target words in the test, the participants had to choose from the scale presented below:

I. I don't remember having seen this word before.
II. I have seen this word before, but I don't know what it means.
III. I have seen this word before, and I think it means _____. (synonym or translation)
IV. I know this word. It means _____. (synonym or translation)
V. I can use this word in a sentence: _____.
(If you do this section, please also do Section 4.)

A participant might have chosen category I and obtain one point in the pretest for he or she had not seen this word before. Then during the summary-writing training, he or she encountered this word in the articles. When taking the posttest, even when the participant might not know the meaning of the word, he or she was still very likely to choose category II (I have seen this word before, but I don't know what it means). However, when it was the case, the participant would obtain two points and thus gain one point in this word compared to his or her pretest performance. The researcher hence calculated the proportion of the two groups' increase in scores that was attributed to this phenomenon. The average improving score gained from this practice effect was calculated and the result is presented in Table 23. The high-level reading group gained 25.21 in the Vocabulary Knowledge posttest compared to the pretest. Their average gained score caused by the practice effect was 6.88. The low-level reading group gained 22.94 in the posttest. Their average gained score caused by practice effect was 7.00. The analysis showed that only a certain proportion of the two groups' vocabulary knowledge improvement was attributed to the practice effect. That is, to a significant extent, both groups had improved their knowledge of the vocabulary selected from the ten articles they had

summarized in the posttest.

Table 23

The Vocabulary Knowledge Posttest Score Possibly Gained by Practice Effect

Group	Test	Mean	Gained score	Average Gained Score Caused by Practice Effect
High-level	Pretest	121.65	25.21	6.88
	Posttest	146.86		
Low-level	Pretest	109.75	22.94	7.00
	Posttest	132.69		

The result is similar to numerous studies which claim that reading can enhance vocabulary learning (Brown, Waring & Donkaewbua, 2008; Day, Omura & Hiramatsu, 1991; Nagy, 1995; Rodriguez & Sadoski, 2000; Pitts, White & Krashen, 1989). Through the summarization process, the participants read the articles, which provided the context of the target words, and the context should have helped improve their knowledge of the target words though the gains might not be great. This finding hence is in line with the conclusion of the previous studies that reading provides learners the context in which the acquisition of vocabulary knowledge can be enhanced (Miller, 1995; Nagy, 1995; Rodriguez & Sadoski, 2000; Webb, 2008).

Despite the significant improvement in vocabulary knowledge, the high-level reading group did not have significant improvement in reading comprehension. This might be because the Vocabulary Knowledge Test was an achievement test, which had been designed to measure the test-takers' ability within a limited range of knowledge, while the reading-comprehension tests in this study were proficiency tests, which had aimed to measure the test-takers' general reading-comprehension proficiency. To demonstrate improvement in a limited range of knowledge is less difficult than to improve language proficiency. This can possibly explain why the high-level reading group demonstrated a significant improvement in vocabulary knowledge but not in reading comprehension. Another explanation might be the possible practice effect of Vocabulary Knowledge Test. Unlike the reading comprehension pretest and

posttest, which were designed to measure reading proficiency and were composed of different test content, the Vocabulary Knowledge pretest and posttest consisted of the same 52 words selected from the passages used for summarization. The practice effect of the same test may have more or less contributed to the participants' progress in the Vocabulary Knowledge posttest.

4.2.3 Summary-writing

In the "Copying" aspect, both high-level and low-level reading groups had demonstrated lower percentages of copying behavior upon the completion of the summary-writing training. In other words, through the WriteToLearn training session both groups made some progress in using their own words when writing a summary. The participants made improvements in copying behavior probably because rewriting the original sentences in their own words is not a skill too complex to pick up. It requires mostly word knowledge and does not seem to be too difficult for the participants as they could use the on-line dictionary to find words similar to the target word in meaning.

Moreover, during the training session, the high-level reading group had higher percentages of copying behavior than the low-level reading group most of the time (Figure 7). In other words, the high-level reading group copied from the original text more often than the other group. To avoid copying words or sentences from the original text, the participants need to digest the content of the articles and manage to express the same content in their own words. The lower "Copying" percentages of the low-level reading group imply that the participants in this group have made more efforts in summarizing the articles than the high-level reading group.

Despite the progress that both groups made on copying behavior, there seems to be no other obvious improvement found in the other three aspects— spelling mistakes, repeated content and unimportant content. As reported previously, both groups generally had very low percentages of spelling mistakes and repeated content from the beginning of the WriteToLearn

training to the end. The percentages of spelling mistakes remained between 0.5 to 2 percent, and those of repeated content were between 0 to 4 percent. However, both high-level and low-level reading groups got a slightly increasing percentage in spelling mistakes during the training session. A possible reason why both groups of participants made a bit more spelling mistakes during the training is that during the summary-writing training, the importance of spelling might not have been emphasized enough. Different from the other aspects, "Spelling" is the only one that does not seem to be a specific requirement for writing a summary, though it seems to be a basic requirement for all types of writing.

On the other hand, the constantly low percentages of "repeated content" in the two groups might be due to the generally limited length of the participants' summaries. Being in the learning stage of writing summaries, the EFL learners generally wrote short summaries and simultaneously were struggling to find the important content. Thus it might be less possible for them to write repetitive content in their summaries.

Finally, the two groups of participants got the highest percentages in the "Unimportant" aspect among the four aspects. Both the "Unimportant" lines of high-level and low-level reading groups fluctuated between 10 and 45 percent during the training session, which was the most unsteady development among the 4 aspects. Moreover, the shapes of the "Unimportant" lines of high-level and low-level reading groups are quite similar. Two factors may have resulted in the high similarity of the two lines. First, although the researcher has tried to control the difficulty level of the ten articles by considering the length of the passages (600 to 1000 words) and the content familiarity to the learners, passages on different topics may have still posed different difficulty level for the participants. The two groups got lower percentages of unimportant content in Summary 3(smoking), Summary 6 (Heroic animals), Summary 8 (Ancient Greece) and Summary 10 (The nation of immigrants), three of which (summary 6, 8, 9) are science-related articles. And both groups of participants got higher percentages in Summary 4 (Online netiquette), Summary 7 (Pandas) and Summary 9 (Earthquake Safety), two

of which are social-science related articles (Summary 7 and 9). It seems that students performed better in grabbing the important content when summarizing social-science related articles than science-related articles. This may be because the participants in this study were foreign-language majors, who were more familiar with social-science related topics and less familiar with science-related topics.

The other possible reason for the unsteady "Unimportant" lines may be the limited capability of the WriteToLearn program to distinguish important content from irrelevant details. During the WriteToLearn training, while maintaining the functions of the program and monitoring the participants' progresses, the researcher found that the WriteToLearn system occasionally misjudged the quality of summaries in the "Unimportant" aspect. The sentences marked by the system as unimportant content could be actually important points based on the researcher's judgment. For instance, in Summary 7, the participants were required to summarize the article "Pandas". The main points of the original article "Pandas" are outlined as follows:

I Introducing pandas

- A The giant panda is one of the best-known and the rarest animal in the world.
- B Pandas live in misty forests in Sichuan since they prefer to be cool and damp.
- C Pandas eat bamboos.

II Why are pandas so rare

- A Pandas do not produce many babies.
- B Pandas need plenty of a certain kinds of bamboo, but the bamboo is not common.
- C Pandas' living environment is threatened by people with the construction of modern transportation.
- D Pandas are still being captured for their skins.

III Saving the panda

A There are thirteen panda reserves in China.

B People are taught the uniqueness of Panda.

C Related fields of professionals and workers are brought together to protect the giant panda.

Below is a summary written by one of the participants. There were some grammatical errors in the summary but as a whole the writer captured the main idea of the original article.

> The giant panda is one of the best-known animals in the world, also one of the rarest. Because it is so popular and unusual, it was chosen as the symbol of the WWF. Pandas prefer to be cool and damp. Dry weather is bad for them. They are also famous for eating bamboo. There are some reasons that explain why pandas are so rare. First, pandas do not produce many babies. Second, they need plenty of bamboo which has a very unusual way of growing. Third, people have captured them from the wild. Although pandas are in danger from hunting, the main threat to them is the loss of their bamboo forests. Besides protecting the pandas, it is important to tell the local people why these animals are so special. If we can give enough support to the efforts to save pandas, perhaps in the next century there will be more pandas in the world than now.

The system reported that in this summary there was 58% unimportant content, which was marked in red. However, the underlined red sentences are considered important content of the passage by the researcher. As can be seen in the outline of this article, the sentence "*Pandas prefer to be cool and damp*" presented the second main point of the first section in this article. The sentences "*There are some reasons that explain why pandas are so rare. First, pandas do not produce many babies. Second, they need plenty of bamboo which has a very unusual way of growing. Third, people have captured them from the wild.*" accurately described the main points of section two, which discussed why pandas are so rare.

Therefore, the system did not always give fair judgment in the importance of the content

in a summary, which made less clear the interpretation of the participants' summary-writing performance in excluding the unimportant content from the summary. Another possible explanation for the high percentages in the "Unimportant" aspect is that to distinguish important content from irrelevant details requires a more complicated processing of information. The participants would have to read through the whole article, omit the irrelevant details, and select the important content. Therefore it is not surprising that the participants performed the least well in the "Unimportant" aspect, since it is a skill that requires complex processing and hence is not easy to master.

Different from the result of Friend's study (2001), which found the effectiveness of summary-writing training on college students' ability in judging the importance of text content, the EFL college learners in the present study did not improve in distinguishing important content from irrelevant details when writing summaries. Their performance, as shown in Table 21, was unsteady throughout the training. It might be partially due to the limited capability of the WriteToLearn program in judging important versus unimportant content.

In addition, the present study is different from Friend's study in participants. The participants in Friend's study were freshmen in the United States who spoke different native languages but all spoke fluent English. The participants in the present study were EFL freshmen in Taiwan, who did not live in an English-speaking environment. In other words, the proficiency level of the participants in the present study was lower. They might need more time to develop skills in judging the important content. This may be another reason for the different results in the two studies.

The participants' not demonstrating improvement in judging important content in the present study is also different from the finding of Franzke et al. (2005). Franzke et al. (2005) found that the 8th graders who had written summaries with the help of WriteToLearn performed better in overall summary quality and content coverage in comparison with the control group. The factor contributing to the difference of findings might be the different grading criteria.

Franzke et al. (2005) used human raters to examine the quality of summaries, whereas the present study examined the quality of the summaries on the basis of the feedback generated by the WriteToLearn system. WriteToLearn's limited capability of judging important content, as was found in the present study, might have under-estimated the participants' actual improvement on distinguishing important content from irrelevant details.

4.2.4 Summary

In sum, it has been demonstrated in this chapter the performance of the high-level and low-level reading groups before and after the WriteToLearn training session in reading comprehension, vocabulary knowledge and summary-writing. As for their reading comprehension performance, the high-level reading group outperformed the low-level group on the pretest but not on the posttest. In addition, the high-level group did not show any improvement in reading comprehension upon the completion of the summary-writing training, while the low-level group made a significant progress. This might be because the high-level group did not need to engage themselves so much as the low-level reading group, for more of the high-level reading participants passed the default threshold of section coverage set by WriteToLearn in the first summary drafts than the low-level reading participants.

As for vocabulary knowledge, there was no significant difference between the high-level and low-level reading groups' performances on the Vocabulary Knowledge pretest, nor was there significant difference on the posttest. Both group made significant improvements in Vocabulary Knowledge Test upon the completion of the summary-writing training. This finding hence is in line with the conclusion of the previous studies that reading provides learners the context in which the acquisition of vocabulary knowledge can be enhanced.

Finally, the two groups' summary-writing performances were examined according to the feedback provided by WriteToLearn (i.e. copying behavior, spelling mistakes, repeated content and unimportant content). Upon the completion of WriteToLearn training, both high-level and low-level reading groups somewhat showed slight progresses in the "Copying" aspect, possibly

because rewriting the original sentences in their own words requires mostly word knowledge instead of the macro knowledge of the whole article and thus did not seem to be too difficult for the participants. Moreover, both groups had very low percentages (lower than 3 percent) of spelling mistakes and repeated content throughout the training. As for the "Unimportant" aspect, the two groups did not seem to make obvious progress through the WriteToLearn training and in fact showed quite unsteady performance. The possible explanations for the participants' unsteady performances on the "Unimportant" aspect might be the differences of passages in content difficulties and the limited capability of the WriteToLearn program to distinguish important content from irrelevant details.

CHAPTER 5

CONCLUSION

5.1 Summary of the Study

The purpose of this study was to evaluate the effect of an on-line summary-writing program, WriteToLearn, on college EFL students' performance in reading comprehension, vocabulary knowledge, and summary-writing. Thirty-five EFL freshman English majors in central Taiwan were grouped into high-level reading group and low-level reading group based on their final exam scores of the previous semester and their reading-comprehension pretest scores. Both groups received the WriteToLearn summary-writing training for 14 weeks. Upon the completion of the training, the researcher examined their performances in reading comprehension, vocabulary knowledge, and summary-writing and conducted analyses to see if the two groups had made significant improvements in the three areas and if there was any significant difference between the improvements of the two groups. The major findings are summarized below.

5.2 Summary of the Major Findings

Dependent t-test results showed that the low-level reading group improved significantly in reading comprehension after receiving the WriteToLearn summary-writing training, whereas the high-level reading group did not make significant improvement in reading comprehension. This might be because more of the summaries of high-level reading participants had met the basic requirement of section coverage than those written by low-level reading participants. The high-level reading participants thus did not need to engage so much in the process of rereading the original article and trying to find out the important content as the low-level reading participants. As such, in comparison with the low-level reading group, the high-level reading group did not benefit much from the summary-writing training in terms of their improvement in reading comprehension.

Both groups of participants improved significantly in their vocabulary knowledge of the 52 unknown words selected from the 10 articles they summarized. In line with previous studies (Brown, Waring & Donkaewbua, 2008; Day, Omura & Hiramatsu, 1991; Nagy, 1995; Pitts, White & Krashen, 1989; Rodriguez & Sadoski, 2000), which pointed out that reading can enhance vocabulary learning, this study further investigated the effect of summary-writing training on vocabulary learning and found the combination of reading and summarization also led to significant learning gains in vocabulary knowledge.

Finally, the two groups' summary-writing performance was examined. Four aspects of feedback provided by the WriteToLearn program: copying behavior, spelling mistakes, repeated content and unimportant content, were analyzed. The researcher found similar patterns in the two groups' performance. Throughout the training session, both groups of participants tended to demonstrate few spelling mistakes and low percentage of repeated content in their summaries. On the other hand, both groups made progress to some degree in reducing copying behavior as shown in lower percentage of copying throughout the training. This is possibly because using their own words in the summary requires mostly word knowledge and is not too difficult to learn. Lastly, both groups' performance on excluding unimportant content was quite unsteady. Further examination of the participants' summaries revealed that the unsteadiness was partially due to the WriteToLearn program's limited capability of judging the importance of the content and mostly due to the learners' limited ability of distinguishing important content from unimportant and irrelevant details. The program's limited judging capability makes the interpretation of the data to some degree questionable. The learners' limited ability of distinguishing important from unimportant content implies that summary-writing involves a deeper level of information processing and to master the skill of selecting important and excluding unimportant information may take more time.

Overall, the WriteToLearn summary-writing program has facilitated the low-level reading

group's reading-comprehension performance, but not the high-level reading group's. However, both groups improved significantly in their vocabulary knowledge of previously unknown words after the summary-writing training.

5.3 Pedagogical Implications

The findings of this study shed some light on our practice of reading instruction. The pedagogical implications are presented as follows. As shown in the study, the WriteToLearn summary-writing program seemed to be more effective in promoting relatively weaker EFL readers' reading improvement. The EFL English-major freshmen who demonstrated relatively poor reading-comprehension performance in the pretest showed significant improvement upon the completion of the summary-writing training. Unfortunately, the relatively stronger readers did not benefit much from using the program.

While acknowledging the strengths of the program, the researcher would suggest that teachers adjust the thresholds levels for different aspects of summary for learners at different reading proficiency level in order to motivate learners to the utmost degree. In other words, adjustments can be made so that the threshold is a bit higher so that the high-level reading group are motivated to revise their summaries and hence engage more in the training and benefit more from the program.

Next, teaching vocabulary can be integrated into the teaching of summary-writing. The results of this study indicate that through the summary-writing training, the EFL college learners not only improved their reading-comprehension but also acquired deeper level of vocabulary knowledge of the words. The learning gains were more of a by-product of summary writing. If teaching vocabulary can be integrated to the teaching of summary writing, the learning gains may be even greater.

Finally, the a little improvement of the learners in copying behavior and unsteady development of excluding unimportant content in summaries imply that learners had more difficulties using their own words and distinguishing important versus unimportant details of

the reading passage. Hence, teachers should consider various ways to help learners learn to use their own words in writing a summary and distinguish important points from irrelevant details when teaching summarization.

5.4 Limitations of the Study

The study was delimited in terms of the participants recruited, the grouping of the participants and the program selected to implement. The 33 participants were all Chinese speaking EFL learners majoring in English and studying in a college located in central Taiwan. The findings of this study may not generalize to other EFL populations. Moreover, in this study, the EFL freshmen were divided into high-level reading group and low-level reading group based on their final examination scores from the previous semester and the reading comprehension pretest. The division of the groups was quite arbitrary. Hence the findings on the difference between the two groups in their posttest performance may not generalize to other EFL learners.

In addition, the limited capability of the WriteToLearn program in judging the unimportant content in summaries made the program less reliable in terms of reflecting the EFL learners' ability of distinguishing the important points from irrelevant details. Since this study examined the EFL learners' summary-writing performance by looking at the 4 types of feedback (copying, spelling, repeated and unimportant) generated from the WriteToLearn program, the possible misjudgment of the program may have limited the interpretation of the data.

Lastly, in this study, the high-level reading group consisted of only 17 participants and the low-level reading group consisted of only 16 learners. The relatively small number of participants may also have limited the statistical power of the data analyses in the study.

5.5 Recommendations for Further Research

This study aimed at evaluating the WriteToLearn program in its effects on facilitating improvement in reading comprehension, vocabulary knowledge and summary-writing. Several recommendations are presented to further explore the use of the program. First, different default thresholds of summary quality can be set for learners at different proficiency levels. In this study, the high-level reading group did not gain much benefit of the WriteToLearn program in terms of their development of reading comprehension possibly because the default threshold of summary quality was not difficult for them to achieve. Having achieved the threshold easily, the high-level reading participants might not have engaged themselves in the WriteToLearn training so much as the low-level reading group. Future research may explore this issue by setting different threshold levels and see if the adjustment leads to better results.

Second, in addition to examining the feedback generated by the WriteToLearn program, future research can further examine the learners' summary-writing performance by having human raters to evaluate the quality of summaries. By doing this, a more holistic view of the usefulness of the program could be obtained.

REFERENCES

Ayres, R. (2002). Learner Attitudes Towards the Use of CALL [Electronic version]. *Computer Assisted Language Learning*, *15*(3), 241-249.

Braine, G. (1997). Beyond word processing: Networked computers in ESL writing classes. *Computers and Composition. 14*, 45-58.

Britt, M., & Sommer, J. (2004). Facilitating textual integration with macro-structure focusing tasks [Electronic version]. *Reading Psychology an international quarterly*, *25*(4), 313-339.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental Vocabulary Acquisition from Reading, Reading-While-Listening, and Listening to Stories [Electronic version]. *Reading in a Foreign Language*, *20*(2), 136-163.

Caccamise, D., & Snyder, L. (2005). Theory and pedagogical practices of text comprehension [Electronic version]. *Topics in Language Disorders, 25*(1), 5-20.

Cobb, K. J. (2002). Facilitating second language acquisition through computer assisted language learning [Electronic version]. *Action Research Exchange (ARE), 1*(1).

Davies, G.. (2000). CALL (Computer Assisted Language Learning). (pp. 90-93). Routledge. Retrieved April 27, 2009, from Education Research Complete database.

Dodigovic, M. (2002). Developing Writing Skills with a Cyber-Coach [Electronic version]. *Computer Assisted Language Learning*, *15*(1), 9-25.

Day, R. R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading [Electronic version]. *Reading in a Foreign Language, 7*(2), 541-551

Foltz, P. W., Gilliam, S., & Kendall, S. (2000). Supporting Content-Based Feedback in On-Line Writing Evaluation with LSA [Electronic version]. *Interactive* Learning *Environments*, *8*(2), 111-127.

Franzke, M., Kintsch, E., Caccamise, D., Johnson, N., & Dooley, S. (2005). Summary Street: Computer support for comprehension and writing [Electronic version]. *Journal of Educational Computing Research*, *33*, 53–80.

Friend, R. (2001). Effects of Strategy Instruction on Summary Writing of College Students [Electronic version]. *Contemporary Educational Psychology*, *26*(1), 3-24.

Garner, R., Gillingham, M. G.. & White, C. S. (1989) Effects of `seductive details' on macroprocessing and microprocessing in adults and children [Electronic version]. *Cognition and Instruction*, *6*, 41-57.

Gajria, M., & Salvia, J. (1992). The Effects of Summarization Instruction on Text Comprehension of Students with Learning Disabilities [Electronic version]. *Exceptional Children*, *58*(6), 508-16.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice.* New York: Cambridge University Press.

Hudson, T. (2007). *Teaching second language reading.* New York: Oxford University Press.

Jitendra, A., Cole, C., Hoppes, M. & Wilson, B. (1998). Effects of a direct instruction main idea summarization program and self-monitoring on reading comprehension of middle school students with learning disabilities [Electronic version]. *Reading and Writing Quarterly, 14*(4), 379-396.

Jitendra, A., Hoppes, M. & Xin, Y. (2000). Enhancing main idea comprehension for students with learning problems: The role of a summarization strategy and self-monitoring instruction. *The Journal of Special Education, 34*(3), 127-139.

Joshi, R, M. (2005). Vocabulary: A critical component of comprehension [Electronic version]. *Reading & Writing Quarterly*, *21*, 209-219.

Kintsch, E., Steinhart, D., Stahl, G., LSA Research, G., Matthews, C., & Lamb, R. (2000).

Developing Summarization Skills through the Use of LSA-Based Feedback [Electronic version]. *Interactive Learning Environments*, *8*(2), 87-109.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production [Electronic version]. *Psychological Review*, *85*, 363-394.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model [Electronic version]. *Psychological Review, 95,* 163-182.

Kintsch, W. (2005). An Overview of Top-Down and Bottom-Up Effects in Comprehension: The CI Perspective [Electronic version]. *Discourse Processes*, 39(2/3), 125-128.

Koda, K. (2004). *Insights into second language reading*. New York: Cambridge University Press.

Landauer, T., Lochbaum, K., & Dooley, S. (2009). A New Formative Assessment Technology for Reading and Writing [Electronic version]. *Theory Into Practice*, 48(1), 44-52.

Landi, N. (2010). An Examination of the Relationship between Reading Comprehension, Higher-Level and Lower-Level Reading Sub-Skills in Adults [Electronic version]. *Reading and Writing: An Interdisciplinary Journal*, *23*(6), 701-717.

Lee, T. (2007). Improving English reading and listening by integrating a Web-based CALL system into classroom instruction [Electronic version]. *Journal of Instruction Delivery Systems*, *21*(3), 21-29.

Lervåa, A., & Aukrust, V. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners [Electronic version]. *Journal of Child Psychology & Psychiatry*, 51(5), 612-620.

Loucky, J. (2003). *Enhancing Japanese college students' English reading and vocabulary skills by using CALL innovations*. IN: Bulletin of Seinan Jogakuin College. (ERIC Document Reproduction Service No. ED482896)

McConkie, G. W., Zola, D., Illinois Univ., U. g., & Bolt, B. A. (1980). Language Constraints and the Functional Stimulus in Reading. (Report No. 194). Urbana, IL: Center for the Study of Reading. (ERIC Document Reproduction Service No. ED 199665)

Miller, S. (1995). Vocabulary Development and Context Usage. Retrieved from ERIC database.

Nagy, W. (1995). On the role of context in first- and second- language vocabulary learning. (Report No.627). Urbana, IL: Center for the Study of Reading. (ERIC Document Reproduction Service No. ED391152)

Helge Niska. (1999). Text Linguistic Models for the Study of Simultaneous Interpreting (chap 5). Stockholm: Department of Finnish, Stockholm University. Retrieved March 15, 2011, from http://www.geocities.com/~tolk/lic/LIC990329.htm

Oded, B., & Walters, J. (2001). Deeper processing for better EFL reading comprehension. *System 29*(3), 357-370.

Paribakht, T., & Wesche, M. (1993a). The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal*, 11, 9-29.

Pitts, M., H. White, & S. Krashen (1989) Acquiring second language vocabulary through reading. Reading in a Foreign Language. *5*(2), 271-275.

Radmacher, S. A., & Latosi-Sawin, E. (1995). Summary writing: A tool to improve student comprehension and writing in psychology [Electronic version]. *Teaching of Psychology*, 22(2).

Rayner, K. & Pollatsek. A. (1989). *The Psychology of Reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rogevich, M. & Perin, D. (2008). Effects on science summarization of a reading comprehension intervention for adolescents with behavior and attention disorders.

Exceptional Children, 74(2), 135-154.

Rodriguez, M., & Sadoski, M. (2000). Effects of Rote, Context, Keyword, and Context/Keyword Methods on Retention of Vocabulary in EFL Classrooms [Electronic version]. *Language Learning*, 50(2), 385-412.

Smith, F. (1994). *Understanding Reading: A Psycholinguistic Analysis of Reading and Learning to Read.* (5th ed.). Hillsdale, N.J. Lawrence Erlbaum Associates.

Son, J. (2007). Learner experiences in web-based language learning [Electronic version]. *Computer Assisted Language Learning*, *20*(1), 21-36.

Sullivan, N. & Pratt, E. (1996). A comparative study of two ESL writing environments: A computer-assisted classroom and a traditional oral classroom. *System,* 29, 491-501.

Sun, Y. (2007). Learner perceptions of a concordancing tool for academic writing [Electronic version]. *Computer Assisted Language Learning*, *20*(4), 323-343.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Wade-Stein, D., & Kintsch, E. (2004). Summary Street: Interactive Computer Support for Writing [Electronic version]. *Cognition and Instruction*. *22*(3). 333-362.

Waring, R. & Takaki, M., (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader [Electronic version]. *Reading in a Foreign Language*, *15*(2), 1-27.

Webb, S., (2008). The effects of context on incidental vocabulary learning [Electronic version]. *Reading in a Foreign Language*, *20*(2), 232-245.

Appendix 1 Sample reading passage selected from the WriteToLearn program

Nutrition Disorders Damage Health

Diets lacking in essential nutrients result in health problems. In addition, the body has a basic Calorie requirement to function. What happens when the body receives too few—or too many—Calories?

Malnutrition

A diet lacking one or more essential nutrients can result in malnutrition. For example, a deficiency in vitamin C can result in scurvy, a disease that causes swollen gums, loose teeth, and small black and blue spots on the skin. Sailors prior to the 1800s often suffered from scurvy because on their long journeys they had no source of fresh fruit and vegetables that provide vitamin C. Today, malnutrition due to protein deficiency is a serious problem in certain parts of the world. For example, higher death rates among children are common in countries where nutrient-rich food is scarce.

Undernutrition

A person whose diet is deficient in Calories is suffering from undernutrition. Eventually the body begins breaking down its own protein molecules for fuel. Muscles shrink and the body may even break down its own tissues to supply energy. Undernutrition usually occurs in parts of the world where drought, war, or some other crisis has disrupted the food supply.

Even simple diets consisting only of rice or corn can provide an adequate number of Calories. However, such diets may be deficient in various nutrients.

Obesity

Consistently ingesting more Calories than are needed can also cause health problems. The human body stores excess Calories as fat until that energy is needed. Health experts generally agree that females can safely have 20-25 percent body fat. Males can safely have 15-19 percent body fat. According to a recent government study, more than 60 percent of Americans are overweight, including 13 percent of children and teens. Obesity is the condition of being seriously overweight, which can have serious negative effects on health, including increased risk of heart disease, diabetes, stroke, and asthma. Most obesity is the result of an imbalance between Calories consumed and amount of exercise. In some cases, genetic and psychological factors play a role.

Eating Disorders

Eating disorders are psychological conditions that affect a person's eating habits and ability to obtain nutrients. Anorexia and bulimia are the two most common disorders. While they do

affect males, these conditions affect a higher percentage of the females in the United States (as many as 4 percent, or more than 7 million women and girls).

Anorexia is an extreme pursuit of thinness characterized by self-starvation and excessive weight loss. A person suffering from anorexia does not consume enough Calories to maintain normal body weight. This condition can lead to disintegration of body organs, menstrual irregularities in women, irregular heartbeat, and even heart failure.

Bulimia is an eating disorder characterized by purging after bingeing (eating excessively). Purging by vomiting or using laxatives results in the premature elimination of food before the body is able to absorb its nutrients. While bulimics do not always appear extremely thin, a bulimic person will often suffer the same health disorders as an anorexic.

A third type of eating disorder is bingeing without purging. As a binge eater consumes excessive Calories, this condition can lead to obesity.

Healthy Eating for Life

Throughout your reading in this chapter you have probably noticed some themes that can help you make healthy decisions about food. One theme is the need for balance between Calories consumed in food and Calories spent in daily activities and exercise. Another theme is the need for a variety of foods in your diet and attention to the amounts of different kinds of foods. One way to get the information you need to make healthy choices is to use the Food Guide Pyramid and to read food labels.

Appendix 2 Sample of Reading Comprehension Test

Article 1

1 It's not just technology that's changed in the last couple of hundred years. Most of us expect that we'll finish our educations and get a job, and that we may change jobs several times throughout our lives. We expect that we'll find a mate, get married and perhaps have children someday, and that those children will grow up and have their own families. Our lives are filled with change - - new places to live, new jobs, new friends. Although it's very ordinary to expect these things, it's also true that someone who lived before the Industrial Revolution might think we were crazy for having such ideas. The Industrial Revolution played a big part in changing lives from predictable ones in rural settings to the more diverse existences we now enjoy.

2 We're used to the idea of constant change, but throughout most of human history, this has not been the norm. For thousands of years, people's lives were much like their parents' had been. A vast majority of the world's population lived in the same village or on the same farm their whole lives. Boys grew up learning their father's work so that they could continue it. Girls' fates were decided by their parents' choice of a husband for them, and there was mostly only one career: wife and mother - -raising the next generation who would again live in the same place, doing the same things.

3 The Industrial Revolution changed many of those patterns. One of the greatest changes was **urbanization**, the move to cities by large numbers of people who lived in rural areas. Beginning in the 19th Century in Great Britain, factories needed huge numbers of workers; these factories were built in cities because **they** needed electricity, gas, water, and roads - - things that were not available in rural areas.

4 People came to take these jobs perhaps not knowing that doing so would change whole societies. No longer would men expect to work their fathers' trades. No longer would women be just another piece of property to be **auctioned** to the man with the most money. Even though most of the work in the factories required little training, it was still training that had to be standardized, so that it could be repeated to lots of workers. This kind of training was one of the foundations of mass education, and it was available - -for the first time in history - -equally to men and women.

5 Industrialization was no picnic, though. Often the work was dangerous and hard, and there were many unscrupulous factory-owners who cheated their workers. The work was often boring and repetitive, and workers were replaceable, so they felt disconnected from their work, partly because factory-work sometimes meant that a worker assembling part of a machine would never see the finished product. Perhaps the most tragic effect of industrialization, though, is pollution. For many years after the beginning of the Industrial Revolution, pollution from factories filled many rivers; more pollution came from the many people who came to work in them. The air became clouded not only from the factories, but also from the products they produced - - machines that created more pollution.

6 If we can dream of deciding to live in almost any city in the world, doing work that our parents cannot even imagine, we get a sense of historical perspective from recognizing that these are ideas that were unheard-of only a couple of hundred years ago, before the Industrial Revolution. From modern-day urban culture, to free public education, to equal rights for women, the effects of the growth of 19th-Century factories truly changed the world in ways the first industrialists could not have foreseen.

_____1. What is the main idea of this passage?
(A) There are many changes in life; people get jobs and have families.
(B) The Industrial Revolution changed the way a lot of people live.
(C) The major effect of the Industrial Revolution is the change in women's status.
(D) 19th-Century industrialists could not have foreseen the changes they would make.

_____2. What does paragraph 2 mainly discuss?
(A) People lived predictable lives for thousands of years in the past.
(B) A couple of hundred years ago, most people were farmers who never went to cities.
(C) Women had much less power in the past and had no choice in the men they married.
(D) Women's main role in life in the past was to bring up the next generation.

_____3. What is the main idea of paragraph 5?
(A) Industrialization did not mean that people got to eat in parks.
(B) Factory-workers during the Industrial Revolution had to work hard.

(C) Air and water-pollution are effects of the Industrial Revolution.
(D) The Industrial Revolution had some bad effects, too.

_____4. Which of the following is **NOT** mentioned as one of the effects of the Industrial Revolution?
(A) Training in factories began and was available equally to men and woman.
(B) Many rivers and the air became polluted by factories in the cities.
(C) Cities grew in size because people moved to the cities to work in the factories.
(D) Girls started to learn their fathers' work so that they could continue it.

_____5. According to the passage, which of the following statements is **NOT** true?
(A) Before the Industrial Revolution most people lived in the same village their whole lives.
(B) Factory-owners during the Industrial Revolution were known to treat workers very well.
(C) Electricity was not available in rural Great Britain at the start of the Industrial Revolution.
(D) Putting together machines was a job performed in factories during the Industrial Revolution.

_____6. What does **urbanization** mean in paragraph 3?
(A) moving to cities
(B) working in factories
(C) getting a new job
(D) making revolution

_____7. What does **they** refer to in paragraph 3?
(A) factories
(B) workers
(C) cities
(D) rural areas

_____8. What can be inferred about Great Britain before the Industrial Revolution?
(A) Public education was not available to both men and women.
(B) Before the Industrial Revolution, there were no cities in Great Britain.

(C) London, the largest city in Great Britain, is very polluted.
(D) Factory-workers are usually unhappy people.

Article 2

1 The idea of a fish being able to generate electricity strong enough to light light bulbs--or even to knock a man down--is almost unbelievable, but several kinds of fish are able to do this for survival. Even more strangely, this curious power operates in different ways for fish belonging to very different families.

2 Perhaps the best known are the electric rays, or torpedoes, most of which live in warm seas. On each side of the head, torpedoes possess a large organ which contains a series of flat electric plates. One side, the negative side, of each plate is supplied with very thin nerves that are connected to a main nerve coming from a special part of the brain. Electricity passes from the upper, positive pole of the organ downwards, or top to bottom, to the negative, lower pole. Generally it is necessary to touch the fish in two places, which then completes the circuit, in order to receive an electric shock.

3 Another famous example is the electric eel. The system of delivering a shock is different from that of the torpedo in that the electric plates run lengthwise and are supplied with nerves from the spinal cord. Consequently, the current passes along the fish lengthwise from head to tail. The electric organs of these fish are really altered muscles, and like all muscles, **they** tire easily, so the fish are not able to produce electricity for very long. Some South American people take advantage of this fact. They drive horses into the river so that the fish discharge their electricity against the horses without hurting the men. When the electric eels are **exhausted**, they can be caught without danger.

4 The power of producing electricity may serve these fish both for defense and feeding. This capability can be clearly seen in two examples. Firstly, if a large enemy attacks, the shock will drive it away. Secondly, it appears that the torpedo and the electric eel use their current most often to stun smaller fish, so that they can catch and eat them more easily.

5 Fish have many different ways of staying alive in the wild. Nature has provided certain groups of fish with a unique and unusual means of

self-protection and survival--electricity!

_____9 . What is the main idea of paragraph 2?

(A) Electric rays are found in tropical oceans and are sometimes called torpedoes.

(B) Electric rays are a kind of fish that can be dangerous if not handled carefully.

(C) Electric rays can only produce a shock if they are touched in two places at the same time.

(D) Electric rays have an organ in their heads that supplies enough current for an electrical shock.

_____10. What is the main idea of paragraph 3?

(A) Electric eels have muscles that produce energy in an unbelievable way.

(B) Electric eels have a muscle-like organ that produces electricity in a head-to-tail direction.

(C) Electric eels do not generate the same strength of electricity as electric rays do.

(D) Electric eels are often caught by people using horses to tire them out.

_____11. Which word best replaces **exhausted** in paragraph 3?

(A) excited (C) tired

(B) shocked (D) hurt

_____12. What word(s) best replaces the pronoun **they** in paragraph 3?

(A) Fish

(B) Head to tail

(C) Organs

(D) Muscles

_____ 13. What is the main idea of the whole passage?

(A) Electricity is a powerful force in nature that can kill animals, as well as, humans.

(B) Fish produce electricity when a current passes through two poles and a circuit.

(C) Fish such as the electric eel and the electric ray generate electricity for survival and defense.

(D) Fish that use electricity for survival are very dangerous and should be avoided by people.

_____14. What can be inferred about electric eels?

(A) They are related to torpedoes.

(B) They are not related to the torpedoes.

(C) Scientists will find a way to use them as a source of energy.

(D) Horses are not seriously injured during their capture in South America.

Article 3

1 We all know that many more people today are right-handed than left-handed. Can one trace this same pattern far back in prehistory? Much of the evidence about right-hand versus left-hand dominance comes from stencils and prints found in rock shelters in Australia and elsewhere, and in many Ice Age caves in France, Spain, and Tasmania. When a left hand has been stenciled, this implies that the artist was right-handed, and vice versa. Even though the paint was often sprayed on by mouth, one can assume that the dominant hand assisted in the operation. One also has to make the assumption that hands were stenciled palm downward—a left hand stenciled palm upward might of course look as if it were a right hand. Of 158 stencils in the French cave of Gargas, 136 have been identified as left, and only 22 as right; right-handedness was therefore heavily predominant.

2 Cave art furnishes other types of evidence of this phenomenon. Most engravings, for example, are best lit from the left, as befits the work of right-handed artists, who generally prefer to have the light source on the left so that the shadow of their hand does not fall on the tip of the engraving tool or brush. In the few cases where an Ice Age figure is **depicted** holding something, it is mostly, though not always, in the right hand.

3 Clues to right-handedness can also be found by other methods. Right-handers tend to have longer, stronger, and more muscular bones on the right side, and Marcellin Boule as long ago as 1911 noted the La Chapel le-aux-Saints Neanderthal skeleton had a right upper arm bone that was noticeably stronger than the left. Similar observations have been made on other Neanderthal skeletons such as La Ferrassie I and Neanderthal itself.

4 Fractures and other cut marks are another source of evidence. Right-handed soldiers tend to be wounded on the left. The skeleton of a 40- or 50-year-old Nabatean warrior, buried 2,000 years ago in the Negev Desert, Israel, had multiple healed fractures to the skull, the left arm, and the ribs.

5 Tools themselves can be revealing. Long-handed Neolithic spoons of

yew wood preserved in Alpine villages dating to 3000 B.C. have survived; the signs of rubbing on their left side indicate that their users were right-handed. The late Ice Age rope found in the French cave of Lascaux consists of fibers spiraling to the right, and was therefore tressed by a righthander.

6 Occasionally one can determine whether stone tools were used in the right hand or the left, and it is even possible to assess how far back this feature can be traced. In stone toolmaking experiments, Nick Toth, a right-hander, held the core (the stone that would become the tool) in his left hand and the hammer stone in his right. As the tool was made, the core was rotated clockwise, and the flakes, removed in sequence, had a little crescent of cortex (the core's outer surface) on the side. Toth's knapping produced 56 percent flakes with the cortex on the right, and 44 percent left-oriented flakes. A left-handed toolmaker would produce the opposite pattern Toth has applied these **criteria** to the similarly made pebble tools from a number of early sites (before 1.5 million years) at Koobi For a, Kenya, probably made by Homo habilis. At seven sites he found that 57 percent of the flakes were right-oriented, and 43 percent left, a pattern almost identical to that produced today.

7 About 90 percent of modern humans are right-handed: we are the only mammal with a preferential use of one hand. The part of the brain responsible for fine control and movement is located in the left cerebral hemisphere, and the findings above suggest that the human brain was already asymmetrical in its structure and function not long after 2 million years ago. Among Neanderthalers of 70,000 – 35,000 years ago, Marcellin Boule noted that the La Chapelle-aux-Saints individual had a left hemisphere slightly bigger than the right, and the same was found for brains of specimens from Neanderthal, Gibraltar, and La Quina.

_____15. Look at the four squares [■] that indicate where the following sentence could be added to the passage. Where would the sentence below best fit?

The stencils of hands found in these shelters and caves allow us to draw conclusions about which hand was dominant.

Paragraph 1：We all know that many more people today are right-handed than left-handed. Can one trace this same pattern far back in prehistory? ■(A) Much of the evidence about right-hand versus

left-hand dominance comes from stencils and prints found in rock shelters in Australia and elsewhere, and in many Ice Age caves in France, Spain, and Tasmania. ■(B) When a left hand has been stenciled, this implies that the artist was right-handed, and vice versa. ■(C) Even though the paint was often sprayed on by mouth, one can assume that the dominant hand assisted in the operation. One also has to make the assumption that hands were stenciled palm downward—a left hand stenciled palm upward might of course look as if it were a right hand. ■(D) Of 158 stencils in the French cave of Gargas, 136 have been identified as left, and only 22 as right; right-handedness was therefore heavily predominant.

_____16. The phrase **depicted** in paragraph 2 is closest in meaning to
(A) identified
(B) revealed
(C) pictured
(D) imagined

_____17. All of the following are mentioned in paragraphs 1 and 2 as evidence of right-handedness in art and artists EXCEPT
(A) the ideal source of lighting for most engravings
(B) the fact that a left hand stenciled palm upward might look like a right hand
(C) the prevalence of outlines of left hands
(D) figures in prehistoric art holding objects with the right hand

_____18. Which of the following statements about fractures and cut marks can be inferred from paragraph 4?
(A) Fractures and cut marks caused by right-handed soldiers tend to occur on the right side of the injured party's body.
(B) The right arm sustains more injuries because, as the dominant arm, it is used more actively.
(C) In most people, the left side of the body is more vulnerable to injury since it is not defended effectively by the dominant arm.
(D) Fractures and cut marks on fossil humans probably occurred after death.

_____19. In paragraph 5, why does the author mention the Ice Age rope found in the French cave of Lascaux?
(A) As an example of an item on which the marks of wear imply that it

was used by a right-handed person

(B) Because tressing is an activity that is easier for a right-handed person than for a left-handed person

(C) Because the cave of Lascaux is the site where researchers have found several prehistoric tools made for right-handed people

(D) As an example of an item whose construction shows that it was right handed made by a right-person

_____20. The word **criteria** in paragraph 6 is closest in meaning to

(A) standards

(B) findings

(C) ideas

(D) techniques

_____21. Directions: An introductory sentence for a brief summary of the passage is provided below. Complete the summary by selecting the THREE answers that express the most important ideas in the passage. Some sentences do not belong in the summary because they express ideas that are not presented in the passage or are minor ideas in the passage.

Summary:

Several categories of evidence indicate that people have always been predominantly right-handed

-
-
-

Answer Choices

(A) Stencils of right-handed figures are characteristic of cave art in France. Spain, and Tasmania.

(B) Signs on the skeletal remains of prehistoric figures, including arm-bone size and injury marks, imply that these are the remains of right-handed people.

(C) Instruments such as spoons, ropes, and pebble tools show signs that indicate they were used or constructed by right-handed people.

(D) The amount of prehistoric art created by right-handed artists indicates that left-handed people were in the minority.

(E) Neanderthal skeletons often have longer finger bones in the right hand, which is evidence that the right hand was stronger.

(F) Nick Toth, a modem right-handed toolmaker. has shown that

prehistoric tools were knapped to fit the right hand.

Article 4

1 The deserts, which already occupy approximately a fourth of the Earth's land surface, have in recent decades been increasing at an alarming pace. The expansion of desertlike conditions into areas where they did not previously exist is called desertification. It has been estimated that an additional one-fourth of the Earth's land surface is **threatened** by this process.

2 Desertification is accomplished primarily through the loss of stabilizing natural vegetation and the subsequent accelerated erosion of the soil by wind and water. In some cases the loose soil is blown completely away, leaving a stony surface. In other cases, the finer particles may be removed, while the sand-sized particles are accumulated to form mobile hills or ridges of sand.

3 Even in the areas that retain a soil cover, the reduction of vegetation typically results in the loss of the soil's ability to absorb substantial quantities of water. The impact of raindrops on the loose soil tends to transfer fine clay particles into the tiniest soil spaces, sealing them and producing a surface that allows very little water penetration. Water absorption is greatly reduced; consequently runoff is increased, resulting in accelerated erosion rates. The gradual drying of the soil caused by its diminished ability to absorb water results in the further loss of vegetation, so that a cycle of progressive surface deterioration is established.

4 In some regions, the increase in desert areas is occurring largely as the result of a trend toward drier climatic conditions. Continued gradual global warming has produced an increase in aridity for some areas over the past few thousand years. The process may be accelerated in subsequent decades if global warming resulting from air pollution seriously increases.

5 There is little doubt, however, that desertification in most areas results primarily from human activities rather than natural processes. The semiarid lands bordering the deserts exist in a delicate ecological balance and are limited in their potential to adjust to increased environmental pressures. Expanding populations are subjecting the land to increasing pressures to provide them with food and fuel. In wet periods, the land may be able to respond to these stresses. During the dry periods that are common phenomena along the desert margins, though, the pressure on the land is often far in excess of its diminished capacity, and desertification results.

6 Four specific activities have been identified as major contributors to the desertification processes: overcultivation, overgrazing, firewood gathering, and overirrigation. The cultivation of crops has expanded into progressively drier regions as population densities have grown. These regions are especially likely to have periods of severe dryness, so that crop failures are common. Since the raising of most crops necessitates the prior removal of the natural vegetation, crop failures leave extensive tracts of land **devoid of** a plant cover and susceptible to wind and water erosion.

7 The raising of livestock is a major economic activity in semiarid lands, where grasses are generally the dominant type of natural vegetation. The consequences of an excessive number of livestock grazing in an area are the reduction of the vegetation cover and the trampling and pulverization of the soil. This is usually followed by the drying of the soil and accelerated erosion.

8 Firewood is the chief fuel used for cooking and heating in many countries. The increased pressures of expanding populations have led to the removal of woody plants so that many cities and towns are surrounded by large areas completely lacking in trees and shrubs. The increasing use of dried animal waste as a substitute fuel has also hurt the soil because this valuable soil conditioner and source of plant nutrients is no longer being returned to the land.

9 The final major human cause of desertification is soil salinization resulting from overirrigation. Excess water from irrigation sinks down into the water table. If no drainage system exists, the water table rises, bringing dissolved salts to the surface. The water evaporates and the salts are left behind, creating a white crustal layer that prevents air and water from reaching the underlying soil.

10 The extreme seriousness of desertification results from the vast areas of land and the tremendous numbers of people affected, as well as from the great difficulty of reversing or even slowing the process. Once the soil has been removed by erosion, only the passage of centuries or millennia will enable new soil to form. In areas where considerable soil still remains, though, a rigorously enforced program of land protection and cover-crop planting may make it possible to reverse the present deterioration of the surface.

_____22. The word **threatened** in the passage is closest in meaning to
(A) Restricted

(B) Endangered
(C) Prevented
(D) Rejected

_____23. According to paragraph 3, the loss of natural vegetation has which of the following consequences for soil?
(A) Increased stony content
(B) Reduced water absorption
(C) Increased numbers of spaces in the soil
(D) Reduced water runoff

_____24. The phrase **devoid of** in the passage is closest in meaning to
(A) Consisting of
(B) Hidden by
(C) Except for
(D) Lacking in

_____25. All of the following are mentioned in paragraph 9 as contributing to desertification EXCEPT
(A) Soil erosion
(B) Global warming
(C) Insufficient irrigation
(D) The raising of livestock

_____26. It can be inferred from passage 10 that the author most likely believes which of the following about the future of desertification?
(A) Governments will act quickly to control further desertification.
(B) The factors influencing desertification occur in cycles and will change in the future.
(C) Desertification will continue to increase.
(D) Desertification will soon occur in all areas of the world.

_____27. Look at the four squares [■] below that indicate where the following sentence can be added to the passage. Where would the sentence best fit?

This economic reliance on livestock in certain regions makes large tracts of land susceptible to overgrazing.

Paragraph 7: ■(A) The raising of livestock is a major economic activity in semiarid lands, where grasses are generally the dominant

type of natural vegetation. ■(B) The consequences of an excessive number of livestock grazing in an area are the reduction of the vegetation cover and the trampling and pulverization of the soil. ■(C) This is usually followed by the drying of the soil and accelerated erosion.■(D)

_____28. Directions: An introductory sentence for a brief summary of the passage is provided below. Complete the summary by selecting the THREE answer choices that express the most important ideas in the passage. Some answer choices do not belong in the summary because they express ideas that are not presented in the passage or are minor ideas in the passage.

Summary:

Many factors have contributed to the great increase in desertification in recent decades.

-
-
-

Answer Choices

A. Growing human populations and the agricultural demands that come with such growth have upset the ecological balance in some areas and led to the spread of deserts.

B. As periods of severe dryness have become more common, failures of a number of different crops have increased.

C. Excessive numbers of cattle and the need for firewood for fuel have reduced grasses and trees, leaving the land unprotected and vulnerable.

D. Extensive irrigation with poor drainage brings salt to the surface of the soil, a process that reduces water and air absorption.

E. Animal dung enriches the soil by providing nutrients for plant growth.

F. Grasses are generally the dominant type of natural vegetation in semiarid lands.

Appendix 3: Target words on the Vocabulary Knowledge Test

odor	drain	deforestation
withstand	descedent	salvage
obesity	obligate	aftershock
stream	devastation	misty
toxic	nutrient	exhalation
embellish	smuggle	pacifist
assembly	inhale	italics
abbreviation	communist	terrain
fertilizer	etiquette	topple
pastime	domesticate	captivity
sustain	ethnic	gibberish
reinforce	defy	composure
excessive	subduction	epicenter
conservation	malnutrition	mandatory
substance	patrol	seismic
rotation	respiratory	vigilance
recruit	convict	
boldface	emoticon	