



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

교육학석사학위논문

The Effects of Text Length and Question
Type on the Reading Comprehension Test
Performance of Korean EFL Readers

지문 길이와 문제 유형이 한국인 EFL 학습자들의
영어 독해 시험 수행에 미치는 영향

2017년 2월

서울대학교 대학원

외국어교육과 영어전공

배 민 령

The Effects of Text Length and Question
Type on the Reading Comprehension Test
Performance of Korean EFL Readers

by
MINRYOUNG BAE

A Thesis Submitted to
the Department of Foreign Language Education
in Partial Fulfillment of the Requirements
for the Degree of Master of Arts in Education

At the
Graduate School of Seoul National University

February 2017

ABSTRACT

The factors affecting English as a foreign language (EFL) learners' reading comprehension have brought up much controversy. Among such factors, the literature on the contribution of text input, especially text length to reading performance, however, is quite limited in offering convincing evidence. Therefore, the present study examines the text input, especially, the influence of text length on Korean EFL readers' text comprehension and their perception focusing on their inference generation process and their use of strategy depending on the two different reading contexts: short or long passages.

A total of 202 Korean EFL learners participated in the present study; 100 of them were college students and 102 were high school freshmen and sophomores. Half of each group took a reading comprehension test consisting of 4 passages (2 short and 2 long) from the CSAT (College Scholastic Ability Test) and half took the other test format consisting of 4 passages (2 long and 2 short), as well. The long version of the test was made from its original restored version in which one or two paragraphs were added to the truncated one. The analysis of the data was then conducted through running a paired samples *t*-test and a repeated measure two-way analysis of variance (ANOVA).

The results show that the influence of text length on readers' performance, question types (multiple-choice and open-ended inference generation questions) is significant. Moreover, it was proved that both college and high school students performed significantly better on open-ended questions which required an

inference generation process when the longer texts were given to the test-takers. Besides the main effects, a significant interaction effect between text length and question type was shown for the college students of higher proficiency level of English.

On the other hand, in terms of their perception of the two different reading passages, the results of the college and high school readers were not consistent. The college students perceived that their reading comprehension benefited from the long passages rather than the short ones. However, for the high school students, they seemed to prefer the short passages over the long ones, but the difference was marginal and did not have any statistical significance. In addition, the participants' responses in stimulated recall interviews revealed some critical information to indicate that the longer text can better assist the EFL readers' comprehension process.

All things considered, these findings have implications for reading comprehension test developers when it comes to deciding on which reading passages to use while setting up to reading comprehension tests, especially if they are high-stake nationwide public examinations, such as the CSAT. Moreover, they shed more light on the importance of the selected factors which can result in better performance of EFL readers.

Key Words: reading comprehension, text length, question type, inference

Student Number: 2015-21849

TABLE OF CONTENTS

ABSTRACT	i
TABLE OF CONTENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
CHAPTER 1. INTRODUCTION	1
1.1. The Background and Purpose of the Study.....	1
1.2. Statement of the Problem.....	5
1.3. Research Questions	10
1.4. Organization of the Thesis	11
CHAPTER 2. LITERATURE REVIEW	12
2.1. Reading Processes: Lower-Level and Higher-Level Processing	12
2.2. Text Comprehension and Inference Generation Processes	15
2.2.1. Processes of Text Comprehension	15
2.2.2. Inference Generation	17
2.3. Effects of Text Characteristics	19
2.4. Text Length	20
2.5. Reading Assessment	27
2.5.1. Multiple-Choice Questions.....	29
2.5.2. Open-Ended Questions.....	31

2.6. Reading and Test-Taking Strategies	33
CHAPTER 3. METHODOLOGY	37
3.1. Participants.....	37
3.1.1. College Students	38
3.1.2. High School Students	40
3.2. Instruments.....	41
3.2.1. The Reading Comprehension Tests.....	41
3.2.2. Self-Assessment Survey	46
3.3. Procedure	46
3.4. Data Analysis	48
3.4.1. Scoring	49
3.4.2. Statistical Analysis	50
CHAPTER 4. RESULTS	52
4.1. Effects of Text Length on Reading Performance	52
4.1.1. Study 1—College Students.....	52
4.1.2. Study 2—High School Students	62
4.2. Effects of Text Length on Readers’ Perception.....	68
4.2.1. Self-Confidence	68
4.2.2. Sensed Easiness	70
4.2.3. Sufficiency of Information	72
4.3. Stimulated Recall Interview.....	74

4.3.1. Reading Process for Multiple-Choice Questions.....	76
4.3.2. Reading Process for Open-Ended Questions.....	79
CHAPTER 5. DISCUSSION.....	82
5.1. Effects of Text Length on Reading Comprehension	82
5.2. Effects of Text Length on Readers' Perception.....	87
5.3. Readers' Strategy Use on Reading Performance.....	90
CHAPTER 6. CONCLUSION	92
6.1. Major Findings	92
6.2. Pedagogical Implications.....	94
6.2.1. Finding Right Reading Texts	94
6.2.2. Selection of Question Types for Reading Tests	95
6.3. Limitations and Suggestions for Future Research	96
REFERENCES	99
APPENDICES	105
국문초록	125

LIST OF TABLES

Table 2.1 Previous Studies Involving the Effect of Text Lengths	25
Table 2.2 Summary of Task Formats Used in Standardized Reading Assessments	31
Table 3.1 Demographics of College Students	39
Table 3.2 Demographics of High School Students.....	41
Table 3.3 Features of Selected Passages for the Present Study	43
Table 3.4 Two Forms of the Test	45
Table 3.5 Form Assignment	47
Table 3.6 Inter-Rater Reliability in Scoring of Open-Ended Questions.....	50
Table 4.1 Results for the College Students' Performance on Two Versions of the Passages by Proficiency Group	53
Table 4.2 Descriptive Statistics for Reading Test Performance in Two Versions of the Test and Two Different Question Types.....	55
Table 4.3 Results of the Repeated Measure Two-Way ANOVA for the Effects of Text Length and Question Types.....	57
Table 4.4 Pairwise Comparisons for the Interaction between Text Length and Question Type for Group H and Group M	59
Table 4.5 Results for the High School Students' Performance on Two Versions of the Passages by Proficiency Group.....	63
Table 4.6 Descriptive Statistics for Reading Test Performance in Two Versions of the Test and Two Different Question Types.....	64

Table 4.7 Results of the Repeated Measure Two-Way ANOVA for the Effects of Text Length and Question Types.....	66
Table 4.8 Results for Self-Confidence in Two Versions of the Passages by Proficiency Group	69
Table 4.9 Results for Sensed Easiness in Two Versions of the Passages by Proficiency Group	71
Table 4.10 Results for Sufficiency of Information in Two Versions of the Passages by Proficiency Group.....	73

LIST OF FIGURES

Figure 4.1 Results of the College Students' Reading Comprehension Performance	56
Figure 4.2 Profile Plots for Pairwise Comparisons of the Interaction Effect.....	60
Figure 4.3 Profile Plots for the Results of the High School Students' Reading Performance	65

CHAPTER 1.

INTRODUCTION

This chapter introduces the research by presenting the motivation and the organization of the thesis. Section 1.1 discusses the background and the purpose of the study. Section 1.2 explains the statement of the problem. Section 1.3 presents the research questions, and the overall organization of the thesis is outlined in Section 1.4.

1.1. The Background and Purpose of the Study

Reading is one of the most crucial skills that one has to acquire for many fundamental reasons: to function properly in one's society, to attain academic success, etc. (Grabe, 2009). With regard to academic success, specifically, reading skills has been noted as the most critical factor among all other language skills (Bernhardt, 1991; O'Reilly & McNamara, 2007). However, to become a fluent reader, especially becoming a fluent second language (L2) reader, is not as easy as it sounds since reading is a complex process which involves multilevel and multicomponent features (Alderson, 2000; Grabe, 2009; Kintsch & Rawson, 2005; van Den Broek, Kendeou, Kremer, Lynch, Butler, White, & Lorch, 2005). A great deal of research has been conducted to understand what text comprehension is and

what the factors are that influence the readers' ability and inability to decode words. These attempts have been mainly done to find out what characteristics proficient readers possess which distinguish them from less-proficient readers. Many cognitive psychologists and reading researchers have defined reading comprehension in a variety of ways. The very core of text comprehension involves a reader's ability to interpret the meaning from the text and "to mentally interconnect different events in the text and to form a coherent representation of what the text is about" (Kendeou, van den Broek, White, & Lynch, 2007, p. 29).

Graesser (2007) noted that among a number of theoretical models of text comprehension for the last two decades, Kintsch's Construction-integration Model (Kintsch, 1988, 1998; Kintsch & Rawson, 2005) is considered to be the most comprehensive one for reading comprehension. According to Kintsch (1994), "in the situation model, the information provided by the text is elaborated from prior knowledge and integrated into it" (p. 294). He also contends that comprehension is guided by activating the prior knowledge stored in long-term memory from the text in working memory through a bottom-up process (the construction phase), which follows the process that the activated ideas in working memory are integrated into the textual inputs and contents (the integration phase). Kintsch's model assumes that the multiple levels of representation get constructed during a comprehension process and that these principal assumptions are: the surface code, the propositional textbase, and the situation model. In order for readers to have a deeper understanding of the text information, they must not only understand the meaning explicitly expressed in the text, but also construct a "situation model," or

the representation of the situation, which the text describes (van Dijk & Kintsch, 1983).

Reading comprehension can be further explained as including one's ability to make inferences, which is defined by Snow (2002) as the reader's "ability to "construct the text base and the mental models that go beyond the information directly articulated in the text" (p. 108). According to Kintsch and Rawson (2005), inferencing plays a critical role in this process of integration of the information presented in the text and forms a coherent "situation model." Since texts are never fully explicit, in order for the reader to thoroughly understand the meaning, they always have to fill the gaps between different sentences and between information in the text using their background knowledge and the text base information stored in their working memory. For example, when a reader reads the following sentences, "Jenny wanted a new jacket. She started to do extra chores for her family." Even though the phrase "to pay for the jacket," is not explicitly presented, the reader would make a connection and construct the mental model in order to infer the complete meaning.

Reading comprehension can be measured for many different purposes. Among the several different purposes of assessments used to measure reading comprehension, the two major purposes for reading assessment are reading-proficiency assessment (e.g., standardized testing) and assessment for learning primarily (e.g., classroom-based measures). In Korean EFL (English as a foreign language) contexts, standardized reading comprehension tests have been largely employed to assess students' reading abilities. However, when assessing students'

reading comprehension abilities, it is of utmost importance to understand the complexity of the construct of reading. That is, it is the responsibility of test-developers to make sure that the construct of the reading comprehension abilities is clearly identified and defined in order to interpret the performance on the standardized reading comprehension tests as an indicator of the test-taker's reading abilities which we want to measure (Bachman, 1990; Bachman & Palmer, 2010).

Research has shown that a test-taker's performance is affected by a number of diverse factors including the test environment, methods of scoring, question types, topic familiarity, and so on (Alderson, 2000; Bachman, 1990; Bachman & Palmer, 2010). Bachman (1990) discusses that a test-takers' performance is affected by the characteristics of test methods. Among these test method facets, the text length is also a part of the facets of the test input. As all of these in some way or another have the potential to negatively impact the test outcomes, it is the task of researchers and test developers to work toward identifying best practices to make measures more valid. As Bachman and Palmer (2010) noted, since the difficulty of reading tests is determined by an interaction between test items and texts used in the reading comprehension tests, choosing the right texts for assessing the learners' reading abilities should be considered as important (Ozuru, Rowe, O'Reilly, & McNamara, 2008).

There have been a few attempts to investigate the characteristics of the passages used in the standardized reading comprehension tests (Beach, 2008; Commander & Stanwyck, 1997; Jones, 2005; Rothkopf & Billington, 1983; Todd & Kessler, 1971), and specifically, the effect of text length characteristics of the

truncated passages used in reading comprehension tests in EFL contexts (Cha, 1995; Choi, 2011; Hashemi et al., 2014; Mehrpour & Riazi, 2004; Lee, 1999), but these studies were limited to quantitative research alone since they presented only the results of quantitative analysis relying heavily on measurement through multiple-choice items only (see Section 2.3.1. for further discussion with the review of the literature concerning text length). That is, their findings are too restrictive in order to offer convincing evidence of their effects on the readers' reading comprehension processes. Therefore, a closer examination on the reading passages used in the standardized reading comprehension is particularly necessary.

1.2. Statement of the Problem

For the last decades, standardized reading comprehension tests have been the most common forms of measuring test-takers' English reading comprehension abilities (Anderson, Bachman, Perkins & Cohen, 1991) and they have been used extensively in Korean contexts to assess EFL learners' English reading comprehension abilities (Kim & Chon, 2014). In most traditional English reading comprehension tests conducted for Korean EFL learners (e.g., TEPS¹, CSAT², National Assessment of Educational Achievement, etc.), the reading passages tend

¹ Test of English Proficiency developed by Seoul National University

² College Scholastic Ability Test (CSAT) conducted in South Korea as a college entrance exam

to be short in length. They are extracted from a variety of sources including a long journal article, novel, technical research paper or professional literature. The passages may range approximately from 100 to 200 words, and specifically, 148 words average for CSAT (Kim, 2016), and they are often “made deliberately brief to increase the total number of items, which was intended to cover a wide range of topics” (Ozuru et al., 2008, p. 1002), and therefore to reduce the effects of prior knowledge (Alderson, 2000). Proponents of using relatively short, single-paragraph texts in the standardized reading tests such as the TOEFL assert that not only can the variety of the topics be covered in the tests, but also the problem of item interdependence can be prevented by using only one question corresponding to one short passage (Alderson, 2000).

However, some have begun to oppose the practice of using short passages, which had been commonly used in the standardized reading comprehension tests. The researchers in favor of using longer texts argue that the conventional standardized reading comprehension tests which only consist of unrealistically short or contrived passages do not induce reading abilities associated with the real academic situations experienced by the readers in a proper academic context (Flippo & Schumm, 2000; Valencia, Hiebert, & Afflerback, 1994). According to their assertions, short reading passages used in the standardized reading tests do not properly demonstrate the actual reading demands associated with academic coursework (Flippo & Schumm, 2000; Valencia, Hiebert, & Afflerback, 1994). As Perry (1959, as cited in Flippo & Schumm, 2000) noted: “The possession of excellent reading skills as evidenced on conventional reading tests is no guarantee

that a student knows how to read long assignments meaningfully” (p. 199).

Furthermore, in the typical standardized reading comprehension tests, the multiple-choice questions have been widely used for the sake of efficiency and reliability. However, some have argued that the artificial nature of reading comprehension assessments, specifically relying on multiple-choice question formats, is a source of concern with regard to validity of the test (e.g., Daneman & Hannon, 2001; Magliano, Millis, Ozuru, & McNamara, 2007; Katz, Lautenschlager, Blackburn, & Harris, 1990). Critics such as Daneman and Hannon (2001) have argued that the reading comprehension section of the SAT does not measure the test-takers’ comprehension of the passages because they could “perform at better-than-chance levels” (p. 208) even without reading the passages. Furthermore, Katz et al. (1990) have shown that test-takers could do quite well on answering multiple-choice questions even with the passage missing.

In Korean EFL settings, natural oral communication opportunities with native speakers are very limited. The absence of language input in everyday life necessitates high-quality verbal and textual input provided in the English classroom because these can serve as an important influence on their L2 literacy. In other words, the very text dealt with in the language classroom forms the exemplars for the language learners. Therefore, it is of utmost importance that the appropriate texts should be selected when it comes to the English reading. However, relatively little attention has been paid to the issue of choosing the right comprehension texts for learners in Korean EFL settings, assuming that students’ overall reading fluency would increase automatically as they advanced to higher

grades (Lee, 2014; Rasinski, 2012).

Moreover, there are some serious issues when it comes to the English reading materials mainly dealt with in the English reading classrooms in Korea. Since 2004, the Korean CSAT has been based on the passages included in the textbooks published by the government-run Educational Broadcasting System (EBS)³, which are available to everyone, so that the family's socioeconomic status (SES) would no longer be an important factor to determine the success on the exam. According to the Basic Plan for 2016 CSAT released by the Ministry of Education, Science and Technology (MEST), approximately 70% of the content on the test would be directly based on the EBS test-prep materials, which would benefit the majority of the test-takers by setting the standards for the high-stake nationwide public examination. Since they are taking up about 70% coverage of the actual test, the majority of the test-takers are preparing for the CSAT English reading tests by studying and even memorizing the content of the truncated passages in the materials, which are originally designed only for the standardized reading comprehension tests. In these circumstances, it is doubtful whether the language input provided to Korean EFL high school students, which are mainly the single-paragraph texts in the EBS materials, is appropriate enough for them to stimulate an actual reading process similar to authentic academic reading requirements.

Furthermore, standardized reading comprehension tests administered in Korean EFL contexts are quite demanding. They require the readers to make inferences,

³ The EBS English-language textbooks are the test-preparation materials for the CSAT published by the Educational Broadcasting System (EBS).

including the ability to understand the main idea, to interpret the author's main message, or to predict what phrase is the most likely choice to fill into the blank space provided even when the information needed to answer the question is not explicitly available in the passage. However, these increased demands on the test-takers pose serious challenges for some struggling Korean EFL readers. For students who acquired some of the basic decoding skills needed, it is one thing to understand a text at a textbase level, and quite another to make appropriate inferences and interpret the text, which is often required in the high-stake reading comprehension tests. Since higher-order skills, such as inferencing can be much more difficult to learn and master than lower order processes, the difficulty of preparing L2 readers to respond to the high-stake reading comprehension tests in Korean secondary schools is especially daunting.

The purpose of this study, therefore, is to build on the existing literature base to test the effects of the amount of text input, especially, the influence of text length given to readers of standardized reading comprehension tests. The present study also examines the effect of question type on Korean EFL readers' reading comprehension. Thus, the materials in this study consist of two versions of reading passages: One is reading passages taken directly from the reading section of the CSAT. The other version consists of the longer reading passages, which were restored from the original texts. The effects of the two different versions of passages (short and long) are examined using the data taken from, not only quantitative research methods, but also from analyzing the qualitative data taken from stimulated recall interviews. In addition, the two different types of questions

will be investigated by comparing the readers' performance of multiple-choice and open-ended questions in two different reading contexts.

1.3. Research Questions

The present study examines how the length of the passages and the types of questions used for the standardized reading comprehension tests administered in Korea affects L2 reading performance among Korean EFL readers. In addition, the impact of text length on the test-takers' perceptions are also investigated using a self-assessment survey and their reading and test-taking strategy use are also investigated using an in-depth stimulated recall interview. The following three research questions are addressed in the study:

1. Do the text length and the question type have an influence on Korean EFL readers' reading comprehension performance? If so, to what extent do the two variables have an effect on Korean EFL readers' reading comprehension?
2. What is the Korean EFL readers' perception of the two different lengths of the passages?

3. What are the reading and test-taking strategies and generated inferences reported by the Korean EFL readers while reading a short passage and a long passage?

1.4. Organization of the Thesis

The present study consists of six chapters. Chapter 1 introduces the topic of the thesis, explicates the background of the study, and develops into the statement of the problems based on previous research and the current educational situation regarding the Korean EFL contexts. It also presents the research questions. Chapter 2 describes the notions of lower-level processes and higher-level processes and Kintsch's Construction-Integration Theory which serves as the theoretical frameworks for understanding the factors that are related to the ability of students' inference generation processes. The literature on reading assessment is also reviewed. In Chapter 3, the method of the study is described in terms of the participants, the instruments, the procedure, and the data analysis. Chapter 4 presents the results and the research findings are discussed in Chapter 5. Finally, Chapter 6 concludes the research with a summary of the major findings, the implications of the present study, and the suggestions for future research.

CHAPTER 2.

LITERATURE REVIEW

This chapter reviews the theoretical literature related to the current study. Section 2.1 discusses the main component abilities of reading comprehension: lower-level processes and higher-level processes. In section 2.2, a review of the theoretical foundation to the present study follows: Kintsch's Construction-Integration Theory and inference generation processes. In Section 2.3, a review of the textual factors related to the ability of reading comprehension follows: focusing on the concept of text length as a factor which affect EFL learners' reading comprehension. Section 2.4 explores previous studies on reading assessments explicating construct validity in the reading comprehension tests and reading and test-taking strategies.

2.1. Reading Processes: Lower-level and Higher-Level Processing

The notion that reading processes are comprised of lower-level processes and higher-level processes appears to be strongly supported by a large amount of research (Urquhart & Weir, 2014; Grabe, 2009). The following lists are the main component abilities of lower-level processes outlined by Grabe (2009).

These (lower-level) processes include word recognition, syntactic parsing (using grammatical information), and semantic-proposition encoding (building clause-level meaning from word meanings and grammatical information). These processes are carried out as part of one's working memory, the framework in which cognitive processing and knowledge resources are integrated for comprehension (p. 22).

As stated above, lower-level processes include some basic units of a language, such as speech sounds and word-spelling. In terms of lower-level processing, the terms and definitions seem to be uniformly specified in the reading comprehension literature. Furthermore, with regards to lower-level processing in L2 reading abilities, substantial evidence has been provided by extensive previous research: "orthographic processing" (Koda, 1999), "phonological coding" (Koda, 1998), "word recognition" (Akamatsu, 2002; Fukkink, Hulstijn, & Simis, 2005; Perfetti, 1985), "working memory activation" (Yoshida, 2003), and "syntactic parsing" (Droop & Verhoeven, 2003).

Some argue that these lower level comprehension skills are critical for comprehension (Muter, Hulme, Snowling & Stevenson, 2004). Others, however, propose that the lower-level processes are just part of the puzzle that enables the reading comprehension process to be explained. That is, these researchers claim that there are other aspects of the comprehension process in addition to the lower level processes, such as inference and literal comprehension of information in the text (Oakhill & Cain, 2012; Oakhill and Cain, 2007) because discourse comprehension goes beyond word and sentence-level understanding (Hogan, Bridges, Justice, & Cain, 2011), and it is referred to as higher-level processes.

According to Grabe (2009), higher-level processing involves both “building a text model of comprehension (emerging in working memory)” and “a situation model.” The former requires readers to “understand what the text itself is trying to signal” while the latter involves “the reader to combine background knowledge with text information and assists in an effective interpretation of the text in line with the reader’s goals” (p. 39-49). In addition to the two models of comprehension, there are also different component skills that make up higher-level processing including using a set of reading strategies, goals, generation of inferences, activating relevant background knowledge, and comprehension monitoring skills (Grabe, 2009; Oakhill & Cain, 2007; Palinscar & Brown, 1984). Therefore, in order for the reader to build a clear mental representation, it is crucial to construct a text model of comprehension and a situation model with a good command of inference generation skills and reading strategy use.

All in all, these two levels of reading processes form the cognitive processing resources that help us accomplish reading comprehension for various purposes. The focus of the current study is on the higher-level processing and most of all, Korean EFL readers’ inference making processing.

2.2. Text Comprehension and Inference Generation Processes

In this section, the text comprehension process from the perspective of the situation model is reviewed starting from the description of three different levels of the readers' mental representation. Next, the review of Kintsch's Construction-Integration (CI) Model is followed, which is directly related to the present study. Lastly, the processes of inference generation, the focus of the present study, are examined.

2.2.1. Processes of Text Comprehension

Some issues regarding how readers manifest meaning in their encounters with texts have been addressed by a number of cognitive psychologists. Notable among them are van Dijk and Kintsch's (1983). Their model of text comprehension includes three levels of comprehension. First of all, with respect to the "surface" level, the structural features of text, "the words and phrases themselves are encoded", along with "linguistic relations between them". At another level of comprehension, referred to as the "textbase" by the authors (van Dijk and Kintsch, 1983), "readers construct the semantic meaning and rhetorical structure of the information in the text" (Kintsch, 1994, p. 294). This level encompasses the interrelated propositions in a complex text, which is labeled as microstructure, and

the hierarchical organization of these various propositions, which is labeled as macrostructure of the text (Kintsch & Rawson, 2005). However, the readers must construct what van Dijk and Kintsch (1983) have called the “situation model” in order to have a deeper understanding of information presented in the text. During the text comprehension process, the ‘situation model’ is acknowledged as core since discourse comprehension can be completed successfully only when an integration between information gathered from the text and information already received prior to reading the text occurs. This process leads to the creation of a new situational model or furthers the development of an existing structure in the knowledge base of the reader. Therefore, text comprehension can only be achieved by understanding a text, which is referred to as the “textbase” level. At this level of text comprehension, the readers are engaged in task of recalling and summarizing. At the same time, by learning from the text, the readers construct the ‘situation model’ which include tasks such as inference and interpretation of the meaning of the text.

Kintsch’s Construction-Integration (CI) Theory of cognition (Kintsch, 1998) is a framework, which is related to a general model of text comprehension for explaining inference generation in reading comprehension and the relevant processes in which the readers engage when they infer the meaning of the text. According to Kintsch (1988), discourse comprehension comprises two phases: construction and integration. In the construction phase, the reader over generates the incoming information with the command of lower-level processes, such as word recognition, syntactic parsing, and proposition information (Grabe, 2009;

Zwaan & Singer, 2003). Then, in the integration phase, the information is consolidated into coherent representation if it can be added to the existing text model by suppressing any irrelevant or weakly activated information. Therefore, if the reader can construct the meaning of the information being read in the text, then she can “undergo integration, which results in a well-structured mental representation” (Kintsch, 1998, p. 95), which basically means that the text is comprehensible to the reader.

2.2.2. Inference Generation

In Kintsch’s Construction-Integration (CI) Model, inference plays a central role in the process of integration of the information presented in the text and form a coherent “situation model” (Kintsch, 1998). For example, a reader may read the following sentence.

(a) Fred parked the car. He locked the door. (Kintsch & Rawson, 2005, p. 219)

Even though there are no explicit markers that connect the two sentences, the readers automatically realize that the ‘door’ is the ‘car door.’ In this case, the gap is rather local, but sometimes the gap may be global, specifically when the readers themselves are left to construct the general theme of the story, which is not

explicitly stated in the text. In addition, the readers also bring up a piece of relevant knowledge to fill the local gap, such as the fact that cars have doors. This is a typical example of knowledge-based inference. However, there is another example of making inferences. See the following example.

(b) Jack missed his class because he went to play golf. He told his teacher he was sick. (Kintsch, 1998, p. 106)

In order to infer the meaning of these two sentences, the reader would need to make a connection and construct a mental representation. By reading (b), the reader would construct a situation model by elaborating ‘Jack lied.’ The examples illustrate above represent two types of inference generations: “bridging inference” and “elaborative inference” (van Dijk & Kintsch, 1983, p. 49). The bridging inferences is needed to construct the coherence of the text. They provide conceptual links between the ideas presented in the text explicitly, including anaphoric and pronominal inferences. On the other hand, elaborative inferences are built on the basis of the reader’s knowledge of the world related to the concepts and events described in the text (Magliano, Millis, Levinstein, & Boonthum, 2011). Therefore, the inference generation is a necessary part of the text comprehension process when we think of the situation model constructed by the readers.

2.3. Effects of Text Characteristics

Learners engage in many types of reading for different purposes in academic settings. For instance, if some learners read narrative stories for entertainment, then their goal of reading is basically to understand the sequence of events that happened in the story. On the other hand, for some learners who are studying new information from expository texts about social studies or scientific facts, then their purpose would be to fully comprehend the new concepts presented in the texts and internalize the information into their long-term memory. Therefore, the level of the reader's comprehension of the text is affected both by readers' competence in generating proper inferences and their specific purposes for understanding the text (Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992). Furthermore, how readers apply their skills and knowledge is also associated with the characteristics of texts (McNamara & Kintsch, 1996). McNamara and her colleagues' research indicated that there is some interaction between the readers' performance and the characteristics of the texts (e.g., cohesion); low-knowledge readers comprehend high-cohesion texts better and the high-knowledge readers were benefited more from low-cohesion texts than high-cohesion texts (McNamara, Kintsch, Songer, & Kintsch, 1996)

Generally, literature concerning the factors affecting readers' ability to create appropriate inferences classified the variables into two broad categories: reader variables and text variables (Alderson, 2000). In previous literature, different

reader factors have been investigated: readers' comprehension skills, their understanding of the demands of the reading tasks, working memory capacity and their background knowledge related to the topic (Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992). The readers' abilities to comprehend the text can also be affected by text-related factors, such as text content, text genre, text structure, and text readability (Taylor & Samuels, 1983). One study conducted by Kobayashi (2002) has investigated the effects of text organization/text types (from loosely to tightly organized text) and item response format (close, open-ended, summary) on the Japanese EFL readers' reading comprehension. She has found that text organization has a significant impact on EFL students' reading performance as it pertains to three different measures.

2.4. Text Length

As listed above, previous research in relation to the effect of text characteristics on reading comprehension mainly focused on the content analysis or text structure and relatively few studies have examined the length of the text. With regards to the effects of the passage length on reading comprehension, Alderson (2000) maintains that, "a problem all reading test developers face is how long the texts should be on which they base their tests" (p. 108). He goes on to argue that few studies have previously researched the text length as a main variable. However, the

argument that the length of the text may have an impact on readers' comprehension process seems legitimate. As many researchers have discussed (Alderson, 2000; Bachman, 1990; Bachman and Palmer, 2010), the length of the text used in the standardized reading comprehension tests is one of the critical factors that have an impact on the readers' performance when it comes to reading comprehension. According to Bachman (1990), a test-takers' performance in reading comprehension tests is affected by the characteristics of test methods. Among these test method facets of the testing environment, the rubric, the test input and the expected response, "the text length is a part of the facets of the test input affecting the test-takers' test scores" (Bachman, 1990, p. 118).

In the reading comprehension tests administered, a primary issue has become the text input, with a specific focus on the length of passages. Passages that are too long can be unwieldy, burdensome and may end up assessing endurance rather than language. As Chastain (1988) noted, language teachers intuitively think that the longer the passage, the more difficult it will become because the longer passages containing a greater number of ideas could be a cognitive burden on the students. This tendency is due to the natural result of their idea of reading "as a laborious process of deciphering a complex and sometimes unfamiliar linguistic code for which the students were required to know all grammar and vocabulary" (p. 234). This claim is in line with the supporters of using numerous short passages in the standardized reading comprehension tests. They maintained that by using several short passages each corresponding question could minimize the content bias by covering a wider variety of topics and decrease the potential problem of

item dependency (Alderson, 2000; Ozuru, Rowe, O'Reilly, & McNamara, 2008).

On the other hand, others claim that passages that are too short may not provide enough context and lack clues, as is the case with authentic reading. In these cases, even though the reading passages themselves used for these exams don't contain any problems in terms of linguistic level, it could lead the readers to develop serious problems in terms of coherence building of the texts and therefore comprehensibility. One of the earliest researchers who investigated the characteristics of short passages in typical standard reading tests, Sternberg (1991), argues that the length of texts should be a matter of concern when it comes to the cognitive processing need for reading comprehension. According to his claims, "the cognitive processing needed to integrate a longer text is different in kind from the cognitive processing needed to pull together a shorter text (p. 541)" and the shorter passages used in the most standardized reading comprehension tests often fail to measure the higher-level comprehension processing.

Others also contented that due to the brevity of the short text, the structure of the passage often tends to be less apparent than the reading texts read in real-life settings, which leads to lesser chances for building coherence compared to the longer ones (Commander & Stanwyck, 1997; Valencia & Pearson, 1988). This lack of structure often makes readers spend more time and effort reading the text to figure out the mental coherence by utilizing possible reading strategies, such as making inferences about the relationships between texts. They would also read the text repeatedly in order to search for the missing information, or use their background knowledge to integrate those ideas to make the text coherent

(Valencia & Pearson, 1988).

In addition, the claims in favor of using longer texts are in agreement with the authenticity arguments; the reading tasks should reflect the real academic situations where students are often required to read lengthy texts for college-level academic coursework (Alderson, 2000; Crossley et al., 2007; Flippo & Schumm, 2000; Valencia, Hiebert, & Afflerback, 1994). In other words, the texts contained in the tests should be representative of the reading tasks that the students are going to encounter in the real life setting (Magliano, Millis, Ozuru, & McNamara, 2007).

According to their assertions, short reading passages typically used in the standardized reading tests do not attempt to evaluate the actual reading demands of the longer texts which are typically required in most college academic courses. Specifically, Valencia, Hiebert, and Afflerback (1994) noted one of the problems of the majority of standardized reading tests; the standardized reading tests measure only limited skills without capturing the students' higher level literacy abilities, which eventually have negative influences on the curriculum, instruction and learning. Similarly, Flippo and Schumm's (2000) review of the literature concerning the reading tests concluded that none of the standardized reading tests utilized any kind of reading material which actually resemble the typical college reading assignment in length and they also emphasized the need to include "sustained passages of text" (p. 410) in order to design authentic reading tests.

Previous studies concerning various text lengths are categorized according to the research contexts and illustrated in detail in Table 2.1 in order to provide an overview of the relevant trends and patterns among the studies. As outlined below,

previous research has defined text in terms of length in a quite different manner. A review of literature indicates that studies examining the effects of texts varied in length range from forty-four words (Todd & Kessler, 1971) to more than 2000 words (Rothkopf & Billington, 1983). Moreover, the range of the “short” and “long” passages in terms of the length is not consistent among those studies; a “long” passage in one study falls into the category of “short” passage of other studies, and vice versa. However, in general, the researchers have considered any text fewer than approximately 120 words to be short text and any text containing over twice or three times the amount of words than that of the short counterpart to be long text.

Table 2.1 summarizes a number of experimental researches investigated the length effect on the readers’ text comprehension and their perceptions. A careful examination of the previously conducted research findings yields a few interesting points. First, the results in relation to the range of the length between 100 and 300 words did not affect the readers’ performance (Hashemi & Bagheri, 2014; Jalilehvand, 2012; Lee, 1999; Todd & Kessler, 1971). That is to say, the length of the longer passage should be more than twice that of the short passage in order for the readers to consider the two types of the texts differentially. This speculation is in line with the previous research studies, which examined the length effect (Lee, 1999; Yi, 2013). Yi (2013) reviewed the literature concerning the text length in reading tests and noted several relevant issues. One of the issues that she was concerned about was that the experiments dealing with less than 400 words did not yield any statistical significant results.

Table 2.1
Previous Studies Involving the Effect of Text Lengths

Context	Study	Text Length (words)		Ratio (short/long)	Measures	Results
		short	long			
L1	Todd & Kessler (1971)	44, 140, 256		.17 .54	Free recall	No Significant difference
	Rothkopf & Billington (1983)	1056, 1709, 2689		.39 .63	Free recall	Short > Long Significant difference
	Commander & Stanwyck (1997)	260	620	.42	Free recall MCQ Self-rating	Long > Short Significant difference
	Beach (2008)	604	2079	.29	Cloze test	No Significant difference
L2	Hashemi & Bagheri (2014)	254 (average)	351.2 (average)	.72	MCQ	No Significant difference
	Jalilehvand (2012)	218	309	.70	MCQ	No Significant difference
Korean EFL	Cha (1995)	100/ 106	412/ 423	.24 .25	MCQ	Long > Short Significant difference
	Lee (1999)	100/ 120	230/ 270	.43 .44	MCQ	No Significant difference
	Choi (2011)	121.94 (average)	355.5 (average)	.34	Survey Qs	Greater text dependency on short text

Secondly, unlike the studies conducted in English as a first language (L1) and English as a second language (L2) contexts, most of the studies conducted in EFL contexts listed above relied only on multiple-choice tests as a measure of reading comprehension. Specifically, the results demonstrate a pattern of insignificant differences except for a study conducted by Cha (1995). To the contrary, the findings in studies using open-ended question formats (e.g., free recall) and multiple measures of reading comprehension appear to produce significant differences (Rothkopf & Billington, 1983; Commander & Stanwyck, 1997). The notion that the two question formats assess different reading processes can also be found in Ozuru, Brine, Kurby, and McNamara's (2013) research. They concluded that open-ended and multiple-choice format questions assess different aspects of reading comprehension processes. This pattern supports the assertion that multiple-choice questions formats do not provide any insight into actual reading processes the readers go through to arrive at a certain answer (Anderson, 1991; Cohen, 1984, Nevo, 1989). Furthermore, these experimental results can be explained in terms of the authenticity arguments mentioned above; longer texts are more effective tapping at the actual reading demands associated with readers' academic coursework (Flippo and Schumm, 2000; Johnston, 1984; Valencia, Hiebrt, & Afflerback, 1994). As Johnston (1984) noted in this study investigating the effects of prior knowledge in reading comprehension, using longer texts as a reading measurement should be encouraged because "longer texts allow more structure to be built into them and they have greater ecological validity" (p. 237). Even if the longer passage with more authentic text inputs does actually enhance

the text comprehension, there is no way to prove that the readers are involved in the higher-level comprehension of the text unless the measurements of reading comprehension are chosen adequately.

Thirdly, as Yi (2013) noted, the results of the previously discussed research have shown disagreements about the effects of longer passages on the reader's text comprehension process. Although the length of the text in reading comprehension tests has been examined in a few previous studies in EFL contexts (e.g., Cha, 1995; Choi, 2011; Hashemi & Bagheri, 2014; Jalilehvand, 2012; Lee, 1999; Yi, 2013), they have shown mixed results concerning the appropriate length of reading passages to help readers to build a coherent model of text. This fact provides support for the necessity for further studies in this area in order to fill the gap regarding the effect of different text length on the comprehensibility of English text by Korean EFL readers.

2.5. Reading Assessment

L2 readers' reading comprehension abilities that we want to measure are not directly observable. Therefore, the students' overall reading abilities must be inferred on the basis of observed performance, such as their scores on reading-proficiency assessment with standardized testing. In order to interpret the performance on this assessment as an indicator of the test-taker's reading abilities,

which we want to measure, we should first clearly identify and define the construct of the reading comprehension abilities (Bachman & Palmer, 2010). According to Bachman (1990), “the fundamental issue in construct validity is the extent to which we can make inferences about hypothesized abilities on the basis of test performance. In construct validation, therefore, we seek to provide evidence that supports specific inferences about relationships between constructs and test scores” (p. 256). In this sense, when developing standardized reading tests, it is crucial to consider the various types of assessment tasks developed for standardized reading tests in order to collect evidence to support arguments for an L2 reading construct, the specific components of reading abilities, and the creation of the overall reading assessments.

The purpose of using standardized reading assessment in the context of reading is to provide useful information about the readers and help teachers or test-developers to collect valid information to make inferences about the test-takers’ reading abilities for various purposes. Multiple-choice and open-ended questions are two of the most popular question formats used to assess test-takers’ reading proficiency. Even though there has been a general consensus that these two formats assess different reading processes, it has not yet been established clearly how exactly these processes are distinct from each other since there are multiple factors that have an impact on readers’ performance (Ozuru et al., 2013). The goal of the present study is to report meaningful and descriptive analyses of the readers’ comprehension processes for the passages used for the standardized reading tests. Therefore, for the present study, both question formats were used to assess the

readers' comprehension abilities and the different outcomes of the two test formats were compared and investigated.

2.5.1. Multiple-Choice Questions

For the last few decades, standardized reading tests have been the most common form of measuring test-takers' reading comprehension abilities. Among the several techniques to ascertain the language learners' level of reading comprehension, one of the most common testing methods is the multiple-choice format. Since "they are regarded as highly reliable, convenient, efficient and economical in scoring" (Kim & Chon, 2014, p. 62), the multiple-choice format has been greatly used in Korean contexts to assess EFL learners' reading comprehension abilities.

However, this format has frequently been criticized. There has been a long-standing debate over the validity of the multiple-choice question formats used for standardized reading comprehension tests and research has shown that some test-takers could perform above advanced-level even without reading the passages associated with the multiple-choice questions (Katz, Lautenshalger, Blackburn, & Harris, 1990; Daneman & Hannon, 2001). Campbell (1999) also noted that multiple-choice items fail to elicit a higher level of cognitive processing. In addition, some researchers argue that the correct answer may be obtained in a variety of ways. Therefore, they have criticized its lack of insight into the

processes readers undergo to arrive at a certain answer (Anderson, 1991; Cohen, 1984, Nevo, 1989). That is, “the process of reaching the correct answer on a reading comprehension test thus may not reflect the processes involved in actual reading contexts” (Anderson et al., 1991, p. 42). The literature regarding language assessments points out that the assumptions of the test-developers concerning what they test and what they expect the test-takers will go through during the tests often do not correspond with the actual processes that take place (Lee & Ku, 2005; Nevo, 1989; Cohen, 1984; Anderson et al., 1991; Anderson, 1991). If the test developers’ expectations from the respondents do not match the actual processes, which test-takers undergo during testing, this difference would affect the validity of inferences we make on the basis of the test scores. Consequently, if the test-takers do not actually go through the comprehension processes that the test-developers have intended during testing, we cannot say that the test is actually measuring the reading comprehension ability that we intended to measure, which eventually affect the validity of inferences we make on the basis of the test scores.

In order to overcome the weaknesses stated above with regard to using only a multiple-choice question format, in the current study, both of the question types, multiple-choice and open-ended questions are used to assess the readers’ proficiency depending on the two different reading contexts.

2.5.2. Open-Ended Questions

In the EFL Korean context, open-ended questions have been used mainly for educational purposes. For example, every reading section of most English textbooks contains one or more open-ended question at the bottom of each page, and many of the ESL (English as a Second language) reading materials contain quite a number of open-ended questions in order to check the readers' understanding of texts. Some advocates of authentic assessments have proposed the use of open-ended questions as an alternative to traditional multiple-choice questions, not only for classroom-based reading assessments, but for the standardized reading assessments, too. Table 2.2 summarizes the most major task formats, which can be or already have been used in standardized reading assessments in Korean EFL language testing context. These items are included in the summary by Grabe (2009), which are based on his review of the examples presented by Alderson (2000), Hughes (2003) and Weir and Milanovic (2003).

Table 2.2
Summary of Task Formats Used in Standardized Reading Assessments
by Grabe (2009, p. 359)

No.	Question Types
1	Multiple-choice
2	Sentence completion
3	Matching (and multiple matching) techniques
4	Gap-filling formats (rational cloze formats)
5	Short answer
6	Free recall
7	Summary (1 sentence, 2 sentences, 5-6 sentences)

Note. This table has been summarized by the author, based on Grabe (2009)

In the English reading classrooms, teachers utilize questioning strategies to engage students with the deeper reading processes. One of the ways that the teacher can ensure the learners' active engagement in high-level cognitive processing is the strategy, 'ASK to THINK-TEL WHY,' proposed by King (2007). This strategy leads readers to ask each other thought-provoking questions about what they have read, heard, or seen and provides support on how to respond by constructing thoughtful, elaborated responses. The examples of comprehension and thinking questions used in strategy are as followed (King, 2007, p. 272):

Comprehension questions:

Describe ... in your own words.
What does ... mean?
Why is ... important?
What caused...?

Thinking questions (connection questions):

Explain why...
Explain how...
How are ... and ... similar?
What is the difference between ... and ...?
How does ... affect ...?
What are the strengths and weaknesses of ...?
How could ... be used to ...?
What would happen if ...?

The comprehension and thinking questions listed above tap into learners' deeper understanding of the text in order to elicit such cognitive activity, such as self-

explanation, inferencing, speculation, elaboration, and making connections between the text and relevant prior knowledge of the world beyond the text. According to Kintsch's Construction-Integration theories of comprehension (Kintsch, 1988), such cognitive activity is essential to the meaning-making process and promotes coherent building of the text and highly integrated mental representation.

The focus of this research is to assess Korean EFL readers' comprehension process in the two different reading contexts (short and long passages). In order to construct the reading comprehension tests that positively affects the readers' inference generation processes, posing questions which functions as guide for reading, is crucial. Therefore, not resorting only to a multiple-choice format, two response formats—open-ended questions and summary writing—were added to the current study and both multiple-choice and open-ended formats were used to examine the readers' level of comprehension by utilizing open-ended questions to promote their deep understanding of the texts.

2.6. Reading and Test-Taking Strategies

In the contexts of the reading strategy research, there have been long-lasting debates over the appropriate classifications of the language learner strategies. According to Cohen, one of the researchers considered to be an expert in the field

of test-taking strategies, the test-taking strategies employed by the test-takers can be categorized in terms of (a) language learner strategies; (b) test-management strategies; and (c) test-wiseness strategies (Cohen, 2006; Cohen & Upton, 2007).

First, the language learner strategies are defined as “the ways that respondents operationalize their basic skills of listening, speaking, reading and writing, as well as the related skills of vocabulary learning, grammar, and translation” (Cohen, 2006, p. 308). Language learning strategies related to the reading comprehension include “skimming and scanning, memorizing text information, making educated guesses, and looking for definitions, examples, main ideas, or details of meaning in the passages” (Kim & Chon, 2014, p. 63). These reading related strategies can be seen as cognitive processes that the test-takers employ to build the coherence of the text and therefore to be able to find the correct answer to the question.

Next, test management strategies and test-wiseness strategies can both be categorized into metacognitive strategies, which are used to respond correctly to the test items in the test-taking contexts. Test management strategies, especially, share common traits with the metacognitive strategies, such as reading the questions before reading the passage and directly looking for answers from the text, rereading information for clarification, selecting options through eliminating other options, or finding answers in the text through synonyms.

Lastly, according to Cohen (2014), test-wiseness strategies refer to “strategies for using knowledge of test formats and other peripheral information to answer test items without going through the expected linguistic and cognitive processes” (p. 306). For example, in the previously conducted empirical study concerning the

test-taking strategies (Kim & Chon, 2014), the authors revealed that the test-takers utilized the existing clues shown in other options or questions and simply employed the test-wiseness strategies such as choosing the longest option without even knowing the meaning of the sentence.

In a study conducted by Anderson (1991), the individual differences in strategy use by adult L2 learners were examined while engaged in two reading tasks: taking a standardized reading comprehension and reading academic texts. The research investigated individual differences that occurred in strategy use in the two different reading contexts and the characteristics of the good readers that distinguish them from the poor readers. The important result of this study suggests that it is hard to pinpoint a certain set of strategies, which can be used to attribute the success of students reading the two tasks in the study. It was found that both of the good and poor readers seemed to employ the same kinds of reading and test-taking strategies while they were engaged in the two reading tasks. It seems to indicate that it is one thing to know what kinds of strategies there are, but it quite another for a reader to be able to apply them strategically.

There are some of the studies that directly addressed reading and test-taking strategies in Korea for the last 20 years (e.g., Kim & Chon, 2014; Joh, 2006, 2012; Maeng, 2006; Lee, 2002, 2004; Lee & Ku, 2005; Song, 1998). Kim and Chon's (2014) recent study tried to provide information to ascertain whether EFL learners actually undergo the test-taking processes that the test-developers aim to test. Contrary to the previously conducted researcher taking a product-oriented view (e.g., the results of the test scores or English reading proficiency levels), the study

explored the Korean EFL high school students' test-taking processes using questionnaires with proficiency and item type as the grouping variables. The study revealed that there was a great deal of use of well-known reading strategies among the advanced-level high school students, and test management strategies by the lower-level students. The results indicated that simply teaching learners to become test-wise is not sufficient for EFL language teachers. It is also important to encourage learners to work towards building basic and fundamental reading skills consistently. With relation to the reading tasks to elicit the learner's strategy use, most of the previous studies were limited to the tasks of multiple-choice items in the standardized test formats. This lack of diversity in task types could lead to the limitations in explaining the actual reading process that the readers undergo while engaging in the reading tasks and their use of reading strategies.

CHAPTER 3.

METHODOLOGY

This chapter describes the research design and methodology employed to investigate the research questions of the present study. Section 3.1 explains the participants. Section 3.2 provides details on the instruments in terms of reading comprehension tests utilized in the study and the readers' perception on each passage in the tests. The subsequent Section 3.3 discusses the specific data collection procedure and in Section 3.4, the methods for data analysis are described.

3.1. Participants

The purpose of the present study was to determine the effect of text length and question type on the reading comprehension of Korean EFL readers. In order to extend the scope of the study to a broader population of the EFL readers, the study was conducted with two groups of EFL students: college and high school students. After the college students were recruited for the first experiment, additional high school students were included for the second experiment. A total of two hundred and two Korean EFL students participated in the main study; one hundred of them were college students and one hundred and two were high

school freshmen and sophomores.

3.1.1. College Students

One hundred college students who have been studying English as a foreign language for over ten years participated in the present study. Participants were recruited through online bulletin board of a university located in Seoul, Korea. The majors of the participants vary ranging from social sciences to natural sciences: business administration, chemistry, computer information, medicine, and architecture, nursing, and public administration.

A preliminary screening was conducted to control the topic familiarity of the passages used in the tests. Prior to the main study, any college students who were interested in participating were asked to answer the preliminary questionnaire survey. The questionnaire was designed to identify any participants who studied the passages of 2015 and 2016 Korean CSAT English section or who took the Korean CSAT tests within two years. This preliminary screening procedure process allowed the researcher to identify participants who may be familiar with the passages used for the study and to determine if they were qualified for the experiment. In addition, the students with English literature and English education majors were excluded because a number of extraneous factors with relation to those students may muddy the results; they were generally upper-advanced level

English readers with relatively greater test-taking skills, relatively large portion of the students have teaching experience (e.g., private tutoring) and they may be already familiar with the content of the reading passages, and the reading passages used in the present study contain the topic of American literature (e.g., *Walt Whitman*). After completing the survey, the researcher contacted 100 respondents individually, asking them to participate in the main study. Table 3.1 lists the distribution of one hundred college students, who have been divided into three groups according to their proficiency levels on TEPS scores: the upper advanced and advanced group with 800 points or above, the upper intermediate level group with 673 to 792 points, and the lower intermediate level group with 668 points or lower, respectively. For those who don't have TEPS scores, their TOEIC scores have been converted into corresponding TEPS scores. The conversion of the TOEIC scores applied in the present study is followed by the conversion table⁴ presented by the TEPS Council. Table 3.1 describes the distribution of the participants.

Table 3.1
Demographics of College Students

Proficiency level	Range of TEPS scores	Average TEPS score	Gender		Total
			Male	Female	
High	943 - 800	864.40	17	17	34
Middle	792 - 673	747.72	16	17	33
Low	668 - 485	610.12	18	15	33
Total			50	50	100

⁴ The conversion table is provided by the TEPS Council at <http://www.teps.or.kr>.

3.1.2. High School Students

One hundred and two high school students participated in the second experiment. The high school students were recruited from the Incheon area. Among the public high schools in Incheon, two representative groups were selected: Incheon International High School (hereafter IIH) and Dongincheon High School (hereafter DH). Care was taken in the choice of the high school students in order to minimize the potential effects of students' familiarity with the content and topic of the passages. Prior to the main study, the researcher met with the English teachers of both schools to make sure that the students have not studied the passages or any materials similar to them. The learners from IIH were in their freshman and sophomore years ($n = 41$) and information on students' stanine levels⁵ with a mock version of their CSAT indicated that there was almost an equal proportion of learner from higher stanine levels (i.e., level 1 = 29.3%, level 2 = 34.1%, level 3 = 26.8%, and level 4 = 9.8%), and they are considered above average. This should be attributed to the fact that IIH is a Special-Purposed High School. On the other hand, all the learners from DH were in their sophomore year ($n = 61$). Most of the students were at the intermediate and lower stanine

⁵ The practice of using stanine level is a standardized way of categorizing test-takers for the CSAT. Stanines are single digit scores ranging from 1 to 9. The distribution is divided into nine parts (standard nine) where a stanine score of 1 is the highest. The top 4 percent of the test-takers receive a grade of 1, the next 7, 12, 17, 20, 17, 12, 7, and 4 percent receive grades of 2 to 9, respectively (Park, 2014, p.43).

levels (i.e., level 1 = 3.3%, level 2 = 8.2%, level 3 = 14.8%, level 4 = 26.2%, level 5 = 32.8%, level 6 = 9.8%, and level 7 = 4.9%). Table 3.2 lists the distribution of the participants for types of schools, information on student stanine levels with a mock version of their CSAT.

Table 3.2
Demographics of High School Students

N=102

Proficiency group	Stanine levels	Schools		Total
		IIH	DH	
High	1-2	26	7	33
Middle	3-4	15	25	40
Low	5-7	0	29	29
Total		41	61	102

3.2. Instruments

In this section, the instruments used for the present study are described including the reading comprehension tests and a test-takers' self-assessment survey.

3.2.1. The Reading Comprehension Tests

For the present study, four short passages were selected from the reading section of the 2015 and 2016 Korean CSAT. The Korean CSAT is a university entrance

exam in Korea, which is administered by the Korea Institute for Curriculum and Evaluation (KICE). The test is considered to be one of the highest-stake standardized tests measuring the Korean EFL learners' English proficiency. The format of the English reading section is quite consistent each year, which itself sets the national standard for the standardized reading comprehension tests administered in Korea. The test consists of 17 listening comprehension questions and 28 reading comprehension questions, which is 45 in total. Every reading passage varies in length from 63 to 188 words, each followed by one and only one corresponding questions except for the last two passages.

The reading comprehension questions can be grouped into clusters according to the typical forms of reading skill measured: Comprehensive understanding (“main ideas,” and “reordering paragraphs”), factual understanding (“information not true of the passage”), and inferring knowledge (“author’s purpose or claim,” “gap-filling”). Specifically, in the question types measuring the readers’ inferring knowledge, the test requires the reader to interpret the author’s main claim, to identify the author's thesis and purpose, or to predict what phrase is mostly likely fitted in the blank. In the 2015 and 2016 Korean CSAT, there were four questions developed to measure students’ inferring knowledge in the format of gap-filling items. In this question type, students were asked to choose the best options to fill the gap in a given passage.

Since the main purpose of the present study was to investigate the contribution of text length on the students’ reading performance, specifically focusing on their inference generation process in two different reading contexts, the multiple-choice

question measuring the readers' inferring knowledge has been chosen for the analysis. In addition, this item type was considered to be appropriate to investigate the effect of the amount of texts given to the readers by comparing the readers' response to the change in the independent condition (text length). While the readers try to find the most appropriate phrase to fill the blank in the inferring questions, the context of the passage plays an important role (Grabe, 2009). If the readers feel that they do not have enough contexts to guide them to infer the meaning of the sentence surrounding the blank, it is very unlikely that they construct the overall meaning of the passage. Table 3.3 provides the specific information about the reading passages selected for the present study.

Table 3.3
Features of Selected Passages for the present Study

Passage no.	School year (administration)	Genre	Passage type	Word count per passage	Lexile measure
1	2015. 11	business & investing	Short	171	1270L
			Long	367	1090L
2	2016. 11	psychology	Short	160	1190L
			Long	371	1290L
3	2015. 11	social science (mythology)	Short	159	820L
			Long	525	920L
4	2016. 11 (college)	literature	Short	172	1050L
			Long	443	1340L
	(high school)	counseling	Short	132	940L
			Long	308	970L

After the four passages had been selected, each of the passages was made into two different types: short and long. As mentioned in Section 3.1, the short version of the passages was taken from the 2015 and 2016 Korean CSAT. A longer

version of the short passage was extended to two or three times of the short ones in length (i.e., 171, 160, 159, 172, and 132-word passages to 367, 371, 525, 443, and 308-word passages, respectively). The longer version was restored on the basis of the original texts the CSAT passages were taken from.

Since the effects of text length on reading comprehension is the primary focus of the present study, it is necessary to control the effect of the other variables associated with the text length. To ensure the readability of the two versions of the passages, the readability of the two versions of the passages has been calculated using the Lexile readability formula. In addition, a pilot study was administered to examine the validity of the assessments: topic familiarity of the reading passages, text readability, and appropriateness of text difficulty for the students.

For a subsequential process, test items were developed for each text in three formats: multiple-choice questions, open-ended questions (bridging- and elaborative-inference types of question), and summary writing. For the multiple-choice questions, the same questions from the previous Korean CSAT were used. In an attempt to measure different levels of comprehension processes and to induce two different kinds of inferences, the two open-ended question types (bridging inference and elaborative inference type questions) were added. The open-ended questions classified as bridging inference type questions, require readers to establish the relationship between adjacent sentences. For examples, in this type of question, the students are asked to identify what a certain anaphoric expression (e.g., pronouns, noun phrases) refers to in the sentence. The bridging inference type question was thought to be relatively easier than the elaborative

inference type questions since the antecedent of the pronoun or noun phrase is explicitly presented in the text. The elaborative inference generation questions are similar to the bridging inference questions in that they also tap the cognitive processes related to the search of information and integration. However, they induce a slightly higher level of text comprehension because they require readers to integrate the information scattered throughout the text in the larger distances as well as bring up the relevant background knowledge (see Appendix A for the two different versions of the passages and the two types of inference generation questions). Summary writing questions were designed for the purpose of inducing the participants to read the text thoroughly and the scores of the summary writing were not included in the analysis. To eliminate the effect of uncontrolled exterior variables, such as time pressure, the participants were not assigned any time limit during the reading task. All open-ended questions were allowed to be answered in Korean to eliminate any negative effects of using English on their reading comprehension performance. The two forms of the tests and the question types are shown in Table 3.4.

Table 3.4
Two Forms of the Test

Passage no.	Form A	Form B	MCQ	Open-ended Q		Score
				Bridging	Elaborative	
1	Short	Long	1	1	2	4
2	Long	Short	1	1	2	4
3	Short	Long	1	1	2	4
4	Long	Short	1	1	2	4
Total score			4	4	8	16

3.2.2. Self-Assessment Survey

A self-assessment survey was developed to explore the participants' perceptions toward each reading passage, more specifically focusing on their confidence regarding their comprehension of the passages, sensed easiness of the passages, and the sufficiency of information in the passages. The participants were asked to answer the self-assessment survey as soon as they finished answering questions for each reading passage (see Appendix A for the form of the self-assessment survey). Question 1 examined the topic familiarity, which was intended to exclude any participants who already knew the content of the passages and Question 2 was about confidence asking participants how confident they felt that their answers were correct. Question 3 examined the sensed easiness of each reading passage, namely, how easy they thought the given passage was. Question 4 was about sufficiency of information, asking participants whether the text was long enough for them to understand the main idea and find the answers to the questions. A five-point Likert response scale was used for each question.

3.3. Procedure

For the reading comprehension test, two forms, Form A and Form B, were developed to avoid the practice effect and to counterbalance the cognitive load

and the administrative time on each test (Bachman, 1990, p. 183). The form assignment for three proficiency groups are shown in Table 3.5.

The participants were randomly assigned to two groups, with one group taking form A and the other group taking Form B, on the basis of their proficiency level as described in Section 3.1. Two groups were created using stratified random selection which resulted in comparable two samples within each of the proficiency groups. After the participants finished answering the questions of each reading passage, they were asked to answer the self-assessment survey.

Table 3.5
Form Assignment

College students			N=100
Proficiency group	Test format		Total
	A	B	
High	17	17	34
Middle	14	19	33
Low	19	14	33
Total	50	50	100

High school students			N=102
Proficiency group	Test format		Total
	A	B	
High	17	16	33
Middle	20	20	40
Low	15	14	29
Total	52	50	102

After two weeks, the think-aloud procedure was conducted to investigate the readers' cognitive processes qualitatively. Four college students were selected for the stimulated recall interviews, consisting of two participants from each of the high and low group. Each participant was asked to read the same short passage that they didn't get the correct answer to in the previous experiment. Then, they were asked to read the long version of same passage. They were encouraged to report any thoughts that came up while reading the two types of passages and to explain in detail why they chose their answers. In addition, they were asked to elaborate their survey responses and justify their reasons. They were interrupted by the researcher only when any points made were unclear and needed clarification. All of the procedures of the verbal reports were recorded and transcribed by the researcher.

3.4. Data Analysis

This section for data analysis begins by addressing the scoring procedure prior to the statistical analysis. Then, the main statistical analyses are explained in detail.

3.4.1. Scoring

Question response data were scored based on the model answers that were developed by the researcher. To examine the validity of the model answers, two expert judges were invited. One of the judges had MAs in applied linguistics and the other had MAs in English literature and both were in a university doctoral course. Any disagreement was solved through discussion.

The multiple-choice questions were marked by the researcher only since the item type was an objective measure. However, to assess the reliability of marking open-ended questions, one other expert judge was invited (a Korean graduate student studying English education in a university master's course and an experienced high school English teacher in Korea). Any specific information related to the participants was coded and not seen to the raters.

The response data was scored by the researcher, and then the second rater scored 20% of each groups' answers independently: 20 samples from college students and 20 samples from high school students. After marking, the inter-rater reliability between the two judges in scoring the open-ended questions was calculated. The inter-rater reliability is expressed as an intra-class correlation coefficient (ICC)⁶. A high degree of reliability was found between two raters as

⁶ The intra-class correlation assesses the degree to which the measure used is able to differentiate between participants with diverging scores, indicated by two or more raters that reach similar conclusions using a particular tool (Stolarova, Wolf, Rinker, & Brielmann, 2014)

seen in Table 3.6. The average measure intra-class correlation coefficient was .827 with a 95% confidence interval from .822 to .905 for college student samples, and .979 from .971 to .985 for high school student samples.

Table 3.6
Inter-Rater Reliability in Scoring of Open-Ended Questions

	Intraclass correlation	95% Confidence Interval	
		Lower Bound	Upper Bound
College	.870	.822	.905
High school	.979	.971	.985

3.4.1. Statistical Analysis

IBM SPSS Statistics Version 23.0 for Windows was employed for the statistical analysis. For each proficiency group, a repeated measure two-way ANOVA was calculated to examine the effect of text length and question type on test-takers' reading comprehension performance and perception. Dependent variables were test-takers' scores and independent variables were text length and question type. A paired samples *t*-test was conducted to compare means of the students' responses for the self-assessment survey questions after the reading comprehension tests.

In addition to the quantitative analyses, qualitative analyses were also conducted through stimulated recall interview. To investigate the meaning of the results drawn from the quantitative analyses, and to examine their actual reading strategies utilized, two students were recruited from each high and low group.

The data collection and the analysis were conducted based on the methodology described in this chapter, and the following chapter presents the results of the study.

CHAPTER 4.

RESULTS

This chapter presents the quantitative and qualitative results of the study. The effects of text length on Korean EFL readers' performance and their perceptions are reported in Section 4.1 and 4.2 respectively, and Section 4.3 provides an in-depth analysis of stimulated recall interviews.

4.1. Effects of Text Length on Reading Performance

The present study was conducted with two different samples: college and high school students. A total of two hundred and two Korean EFL students participated in the main study. One hundred of them were college students and one hundred and two were high school freshmen and sophomores.

4.1.1. Study 1—College Students

With regard to the first research question (“Does passage length have an influence on Korean EFL readers’ reading comprehension performance? If so, to what extent does the passage length have an effect on readers’ comprehension?”),

the effects of text length on the college students' overall reading comprehension scores were analyzed by conducting a paired samples t-test. The differences of the mean scores between the short and long passages were compared in order to examine the impact of the text length on the readers' performance. Table 4.1 shows the descriptive statistics of overall mean scores and standard deviations for the two versions of the reading comprehension tests and the results of a paired samples t-test.

Table 4.1
Results for the College Students' Performance on Two Versions of the Passages by Proficiency Group

College students								N=100
Proficiency group	N	Text length						p value by t-test
		Short			Long			
		M	S.D.	Range	M	S.D.	Range	
High	34	5.03	1.424	3-8	5.82	1.660	3-8	.040*
Middle	33	4.12	1.515	1-7	5.12	1.556	2-8	.010*
Low	33	4.21	1.536	1-7	4.90	1.627	2-8	.087

Note. Significance level: * $p < .05$

As Table 4.1 illustrates, it was observed that these two types of reading comprehension tests were different in terms of text length for Group High and Group Middle. For the two groups of students (Group High and Middle), the average score was higher for the long passages than short ones (Group High: short

= 5.03 and long = 5.82; Group Middle: short = 4.12 and long = 5.12). For Low group, the students performed marginally better for the long version (short = 4.21 and long = 4.90), but the difference was not statistically significant.

In order to examine the effect of text length and question type on test-takers' reading comprehension performance, a repeated measure two-way analysis of variance (ANOVA) were calculated. The independent variables were the text length (short and long) and the question types (multiple-choice questions and open-ended questions). The dependent variable was the test-takers' reading scores on the test.

For the MCQ items, a total of four questions were given (two short and two long, 1 point was awarded for each correct item), so the highest score and the lowest score one could get for each text type (short or long) was 2 and 0, respectively. For the open-ended questions, there were two items each measuring two different classes of information processing activities: bridging inference and elaborative inference. As for the scoring, 1 point for each bridging inference question, and 2 points for each elaborative inference question were awarded for items correctly responded to, so the maximum score one could get per passage was 3 points, which added up to 6 points for each text type (short or long) in total.

Table 4.2 reports on the three proficiency groups' mean scores and standard deviations for four different sets of reading comprehension tests, and Figure 4.1 illustrates the mean scores for each question type between the two versions of the text.

Table 4.2
Descriptive Statistics for Reading Test Performance in Two Versions of the Test and Two Different Question Types

College students										N=100
		Text length								
Proficiency group	Short				Long					
	MCQ		Open-ended		MCQ		Open-ended			
	M	SD	M	SD	M	SD	M	SD		
High	1.47	.748	3.55	1.078	1.35	.812	4.47	1.186		
Middle	1.15	.712	2.96	1.185	1.12	.819	4.00	1.030		
Low	.93	.704	3.27	1.375	1.09	.765	3.81	1.333		

Note. MCQ: Multiple-choice question (Fill-in-the-blank), Open-ended: open-ended questions (bridging and elaborative inference)

As can be seen from the table above, Group Low obtained higher MCQ scores for the long passages (1.09) than for the short passages (.939) whereas Group High and Group Middle obtained higher scores for the short passages than for the long passages. The mean score of Group High and Middle for the short passage MCQ items were 1.47 and 1.15, which were higher than 1.352 and 1.121, respectively, for the long passage MCQ items. A line chart below depicts changes between the two different text length, showing data and trends.

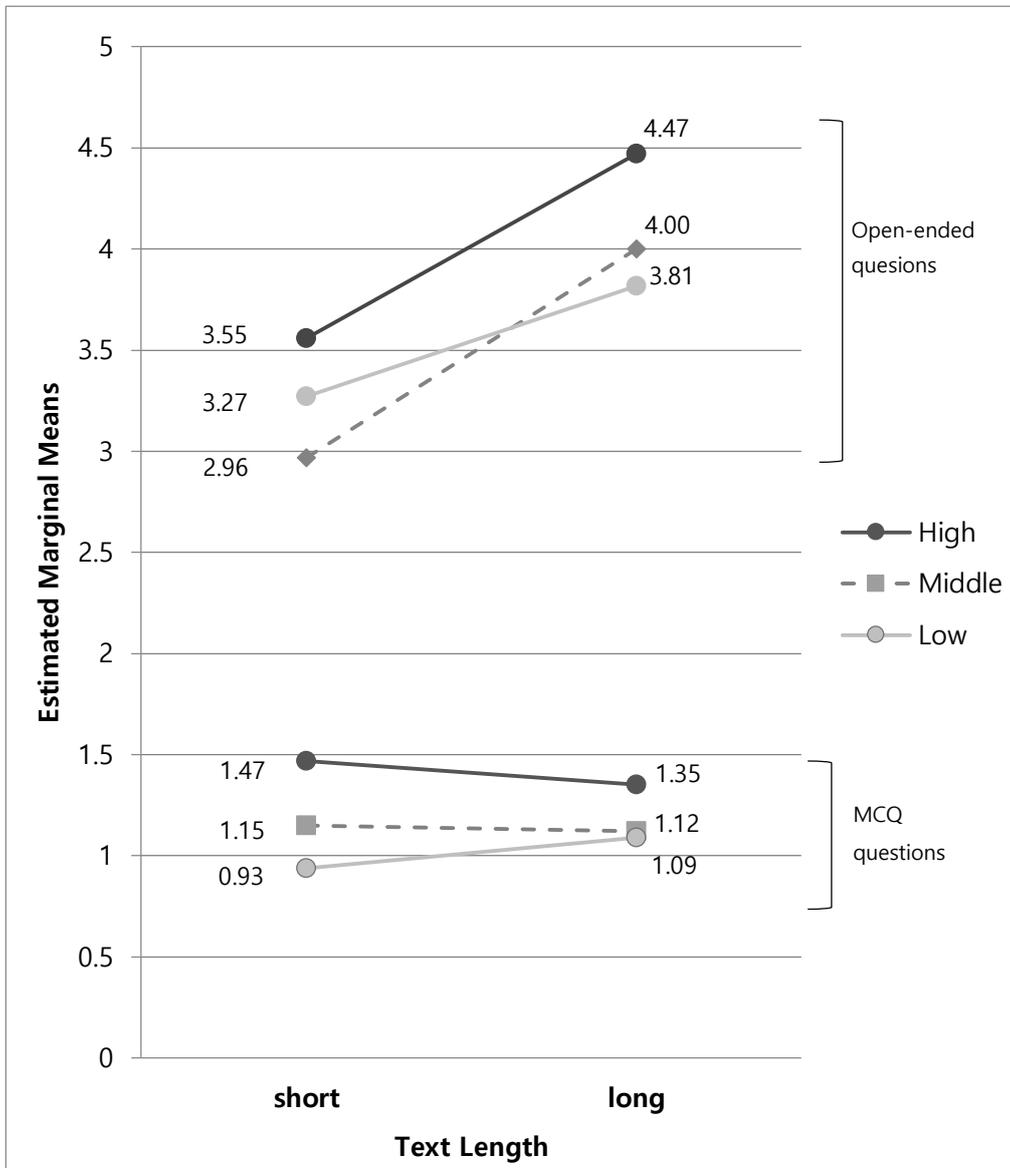


Figure 4.1
Results of the College Students' Reading Comprehension Performance

Figure 4.1 clearly shows that the mean scores obtained for the long version of the open-ended questions are higher than the short one for all three proficiency groups. For example, for the open-ended questions, the mean score of Group High

for the long passages was 4.47, and it was higher than that of the short passages, 3.558. Group Middle's mean score for the long passages was 4.0, which was also higher than its counterpart, short passages, 2.969. However, for the multiple-choice questions, the results were quite different. For the low-proficiency group of college students, the mean score of the multiple-choice questions were higher for long passages than for short passages (short = 0.939 and long = 1.09). For the other two groups of students (Group High and Middle), the score for the multiple-choice questions was not higher for the long passages than short ones. The means were higher with short passages (Group High = 1.47 and Group Middle = 1.15) whereas lower with long passages (Group High = 1.121 and Group Middle = 1.09).

In order to see whether the calculated difference between the means obtained for the short and long version of the test was statistically significant, a repeated measure two-way ANOVA was conducted. The following table presents the results obtained from this statistical analysis.

Table 4.3
Results of the Repeated Measure Two-Way ANOVA for the Effects of Text Length and Question Types

Proficiency group	Source	SS	df	MS	F	p	Partial η^2
High	TL	5.360	1	5.360	4.548*	.040	.121
	QT	230.360	1	230.360	347.281***	.000	.913
	TL*QT	9.007	1	9.007	12.261**	.001	.271
	Error	24.243	33	.735			

Middle	TL	8.250	1	8.250	8.250*	.010	.191
	QT	182.008	1	182.008	358.582***	.000	.918
	TL*QT	9.280	1	9.280	11.435**	.002	.271
	Error	24.970	32	.780			
Low	TL	4.008	1	4.008	3.109	.087	.089
	QT	211.280	1	211.280	147.074***	.000	.821
	TL*QT	1.280	1	1.280	1.578	.218	.047
	Error	25.970	32	.812			

Note. TL: Text Length, QT: Question Type,
Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

Table 4.3 provides a summary of the repeated measure two-way ANOVA. The results showed a significant effect of text length and question type on the college students reading comprehension performance. As for the main effect of text length, significant main effects were found for Group High [$F(1,33) = 4.548, p < .040$] and Group Middle [$F(1,32) = 8.250, p < .010$], and their effect sizes were relatively small: partial $\eta^2 = .121$ for Group High, partial $\eta^2 = .191$ for Group Middle.

In the case of the main effect of question types, a significant main effect was found for all three proficiency groups, Group High [$F(1,33) = 347.281, p < .000$], Group Middle [$F(1,32) = 358.582, p < .000$], and Group Low [$F(1,32) = 147.074, p < .000$]. The effect size of each group were very big: partial $\eta^2 = .913$ for Group High, partial $\eta^2 = .918$ for Group Middle, and partial $\eta^2 = .821$ for Group Low.

Besides the main effects, a significant interaction effect between text length and question type was detected for Group High [$F(1,33) = 12.261, p < .001$] and Group Middle [$F(1,32) = 11.435, p < .002$], and their effect sizes were the same as the partial $\eta^2 = .271$ for both groups. This indicates that text length may affect the college students' reading comprehension performance differently depending on the type of the question. Due to the significant interaction effect between text length and question type in Group H and Group M, it was decided to additionally conduct the test of simple main effects. Simple main effects were computed for each question type across the two different conditions of text length. The following Table 4.4 and Figure 4.3 present the results obtained from the statistical analysis.

Table 4.4
Pairwise Comparisons for the Interaction between Text Length and Question Type for Group H and Group M

Proficiency group	Question type	Text Length		<i>Subjects (F_1)</i>		
		(A)	(B)	<i>Mean Difference (B-A)</i>	<i>Std. Error</i>	<i>p</i>
High	MCQ	short	long	-.118	.178	.513
	Open-ended			.912	.284	.003**
Middle	MCQ	short	long	-.030	.206	.884
	Open-ended			1.030	.266	.001**

Note. Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

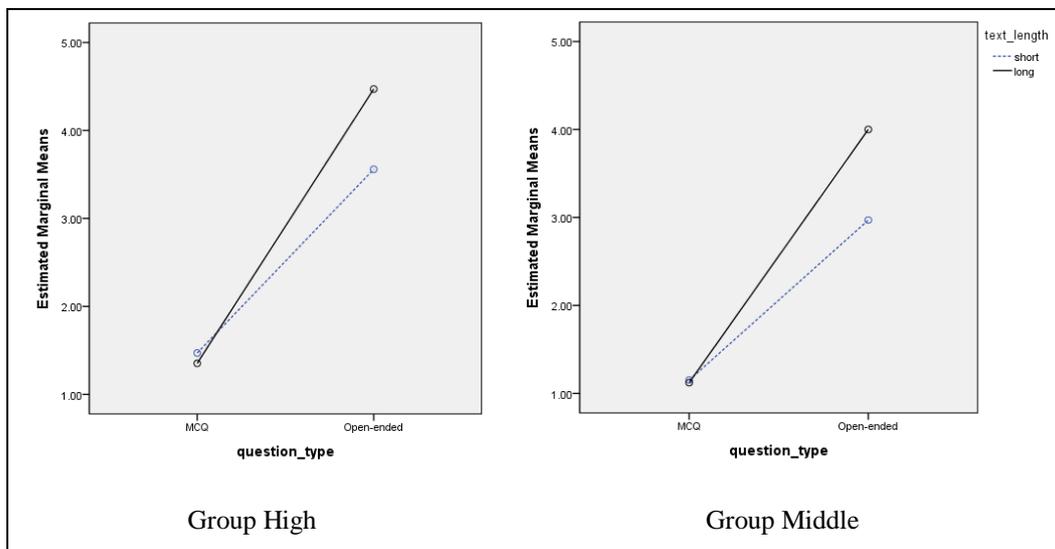


Figure 4.2
Profile Plots for Pairwise Comparisons of the Interaction Effect

As Table 4.4 and Figure 4.3 illustrate, the analysis revealed the significant interaction effects between text length and question type for the Group High and Middle. The mean scores of the multiple-choice questions were not significantly different between the two reading contexts (short and long). For the open-ended questions, on the other hand, the mean scores were significantly different between the short and long passages for both Group High and Middle, $p < .01$. This suggests that the difference between the two reading contexts did not show any statistical significance regarding the multiple-choice questions whereas there was a significant difference between the short and longer text for the open-ended questions. This implies that the longer text had a positive effect on answering the open-ended questions for the readers, but not on the multiple-choice questions.

To sum up, text length was a factor that significantly affected the reading

comprehension of the Korean EFL college students, which ultimately shows the overall superiority of the longer texts over the short texts on the readers' comprehension performance. For the main effect of the text length on the students reading comprehension, the advanced-level and intermediate-level college students, performed significantly better on the long version of the reading comprehension tests. In addition, the question types variable affected the reading performance of all three proficiency groups of the college students. That is, all three groups of the college students obtained higher scores on the open-ended inference questions than the multiple-choice questions. Above all findings, it was especially interesting to note that the interaction effect between the length of text and the type of question has been found among the groups with a relatively advanced level of proficiency. The additional simple main effect analysis revealed that, depending on the question type, the effect of text length might differ. That is, for both advanced and intermediate-level college students, reading the same ideas repeated throughout the passage did assist their comprehension process to find the answers for the open-ended questions. However, reading more text did not seem to be helpful to pick out the best option to fill in the blank space for the multiple-choice question. This might suggest that the readers' comprehension processes might be affected by the type of question they are given. The researcher was curious to know whether this peculiar trend would occur among other Korean EFL readers, specifically, high school learners. In an attempt to answer this question, the second experiment was conducted for high school students.

4.1.2. Study 2—High School Students

In order to gain further evidence for the above-noted findings and to see how the length of the texts would affect the reading performance of other groups of Korean EFL learners, an additional experiment was conducted comprising of 102 high school students and using the same methods as in the first experiment. The high school students' overall reading comprehension scores were analyzed by conducting a paired samples t-test. The differences of the mean score between the short and long passages were compared in order to examine the impact of the text length on the high school readers' performance. Table 4.4 shows the descriptive statistics of average scores and standard deviations for the two versions of the reading comprehension tests and the results of a paired samples t-test.

As Table 4.4 shows, for Group High, the two versions of reading tests were different in terms of text length. For the other two groups of high school students (Group Middle and Group Low), the average score appears to be higher for the long passages than short ones (Group Middle: short = 2.97 and long = 3.25; Group Low: short = 2.37 and long = 2.82). However, the difference between the mean scores of the two groups was not statistically significant.

Table 4.5
Results for the High School Students' Performance on Two Versions of the Passages by Proficiency Group

High School students								N=102
Proficiency group	N	Text length						p value by t-test
		Short			Long			
		M	S.D.	Range	M	S.D.	Range	
High	33	4.72	1.700	1-7	5.45	1.620	3-8	.045*
Middle	40	2.97	1.180	1-5	3.25	1.790	0-6	.470
Low	29	2.37	1.980	0-7	2.82	1.550	0-5	.353

Note. Significance level: * $p < .05$

For the results of the high school students, a repeated measure two-way analysis of variance (ANOVA) were calculated in the same manner as the results of the college students in order to test the effect of text length and question type on the readers' test performance. The independent variables were the text length (short and long) and the question types (multiple-choice questions and open-ended questions). The dependent variable was the test-takers' reading scores on the test.

Table 4.5 below shows the mean scores and standard deviations for four different sets of reading comprehension tests for high school students by three proficiency groups and Figure 4.2 illustrates the between-group differences in the mean scores for each question type between the two versions of the text.

Table 4.6
Descriptive Statistics for Reading Test Performance in Two Versions of the Test and Two Different Question Types

High School students		N=102							
Proficiency group	Text length								
	Short				Long				
	MCQ		Open-ended		MCQ		Open-ended		
	M	SD	M	SD	M	SD	M	SD	
High	1.06	.788	3.66	1.338	1.15	.755	4.30	1.310	
Middle	.47	.640	2.50	1.062	.57	.594	2.67	1.439	
Low	.48	.633	1.89	1.697	.58	.568	2.24	1.299	

Note. MCQ: Multiple-choice question (Fill-in-the-blank), Open-ended: open-ended questions (bridging and elaborative inference)

Table 4.5 presents the descriptive statistics for the two versions of the reading tests and two different questions types for high school students. As depicted in Figure 4.2 above, all three groups of high school students obtained higher scores for the reading comprehension tests when the text given was longer than the shorter one. For the multiple-choice question types, the mean score of Group High was 1.15 for long passages, which was higher than short ones, 1.06. Group Middle's mean score for long text was .575, which was also higher than the short ones, 4.757. In the case of Group Low, the mean score of long passages was .586, which was higher than the short ones, .482.

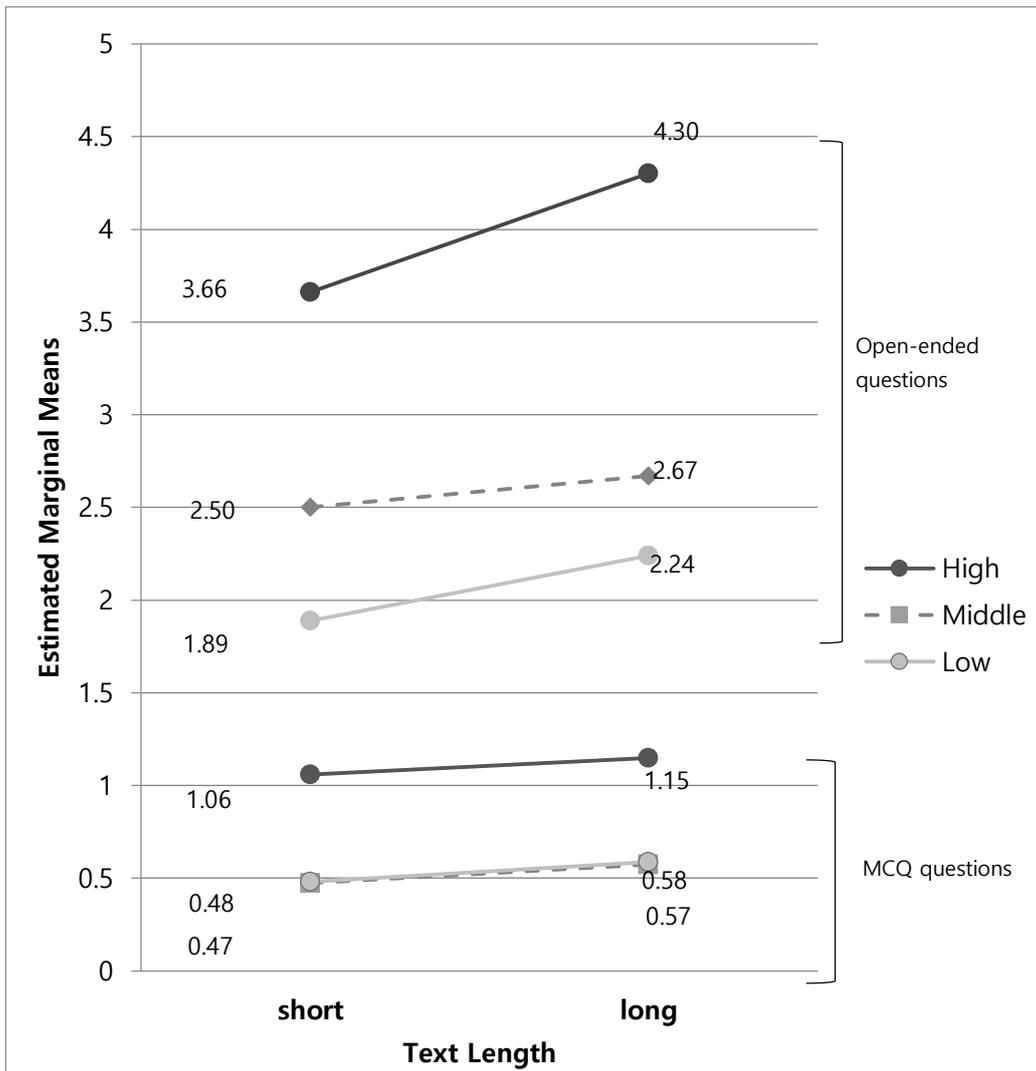


Figure 4.3
Profile Plots for the Results of the High School Students' Reading Performance

In terms of the open-ended questions, the overall trend was the same as the above. All three groups gained higher scores for the long passages than for the short passages (Group High = 4.3, Group Middle = 2.67, and Group Low = 2.24 for long passages, 1.06, .475, and .482 for short passages, respectively). To see if

the marginal differences shown between the means obtained for the short and long version of the test were statistically significant, a repeated measure two-way ANOVA was also employed. The results obtained from the analysis are presented in Table 4.6 below.

Table 4.7
Results of the Repeated Measure Two-Way ANOVA for the Effects of Text Length and Question Types

Proficiency group	Source	SS	df	MS	F	p	Partial η^2
High	TL	4.364	1	4.364	4.345*	.045	.120
	QT	273.485	1	273.485	301.619**	.000	.904
	TL*QT	2.455	1	2.455	2.377	.133	.069
	Error	33.045	32	1.033			
Middle	TL	.756	1	.756	.531	.470	.013
	QT	170.156	1	170.156	220.514***	.000	.850
	TL*QT	.056	1	.056	.064	.801	.002
	Error	34.194	39	.877			
Low	TL	1.457	1	1.457	.891	.353	.031
	QT	68.284	1	68.284	54.681***	.000	.661
	TL*QT	.422	1	.422	.496	.487	.017
	Error	23.828	28	.851			

Note. TL: Text Length, QT: Question Type,
Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

As the table indicated, no significant interaction effect between the text length and questions types was shown in all three proficiency groups of the high school students. Therefore, the main effects of the text length and questions types were

analyzed. For the high proficiency group, there was a significant difference between the performances of the students regarding the two versions of the reading comprehension test [$F(1,32) = 4.345, p < .045$], and its effect size was small (partial $\eta^2 = .120$), which indicates that the students performed significantly better on the long passages than the short ones.

As seen from the results of the college students, a significant main effect of question types was also found for all three proficiency groups, Group High [$F(1,33) = 4.345, p < .045$], Group Middle [$F(1,39) = 220.514, p < .000$], and Group Low [$F(1,28) = 54.681, p < .000$]. The effect size of each group was very big: partial $\eta^2 = .904$ for Group High, partial $\eta^2 = .850$ for Group Middle, and partial $\eta^2 = .661$ for Group Low.

All told, the findings above suggest that the reading comprehension of Korean EFL high school readers is influenced by the text length as well as the question types. Unlike the results from the college students, the interaction effect between the two variables was not detected for all groups of the high school students. Among the three groups of the high school students, high-level high school students benefited from reading the long passages more than the short ones. Even though the difference of the mean scores between the short and long passages did not meet the statistically meaningful level, the trends seen from the results suggest that all levels of high school students benefited from reading the longer texts rather than the short texts.

Findings also suggest that the question types variable affected all three proficiency groups of Korean EFL high school students. As mentioned earlier (see

Section 4.1.2), all three levels of the college students obtained higher scores on the open-ended inference generation questions than the multiple-choice questions and the same trend of difference was also noted for high school EFL readers. That is, all of the proficiency groups of the high school students performed better on the open-ended questions.

4.2. Effects of Text Length on Readers' Perception

To answer the second research question (“What is the Korean EFL readers’ perception of the two different lengths of the passages?”), students’ response survey regarding the confidence, sensed easiness, and sufficiency of information on the two version of the text (short and long) were analyzed.

4.2.1. Self-Confidence

For data analysis, a paired samples *t*-test was conducted to investigate the effect of text length on the EFL readers’ confidence. The dependent variable was scores calculated according to the answers given by the participants. The confidence level in each passage was asked using the phrase “I’m confident that I chose the correct answer.” A five-point Likert response scale was used for each

question (e.g., from ‘5’ very likely to ‘1’ not likely). The results of the descriptive statistics of means and the standard deviation of the students’ confidence and a paired samples *t*-test are shown in Table 4.7 below.

Table 4.8
Results for Self-Confidence in Two Versions of the Passages by Proficiency Group

College students							N=100
Proficiency group	N	Text length				<i>p</i> value by <i>t</i> -test	
		Short		Long			
		M	SD	M	SD		
High	34	2.362	1.448	3.558	.959	.000***	
Middle	33	2.242	1.293	3.09	.971	.000***	
Low	33	1.848	1.1	2.833	.797	.000***	

Note. Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

High school students							N=102
Proficiency group	N	Text length				<i>p</i> value by <i>t</i> -test	
		Short		Long			
		M	SD	M	SD		
High	33	3.303	.991	3.375	.866	.501	
Middle	40	2.312	.951	2.075	.944	.018*	
Low	29	1.724	.830	1.706	.750	.920	

Note. Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

As shown in Table 4.7, in all three proficiency groups of college students, their level of self-confidence was higher for the long passages than for the short passages (Group High: short = 2.362 and long = 3.558; Group Middle: short = 2.242 and long = 3.09; and Group Low: short = 1.848 and long = 2.833). However,

a different trend was found with the high school students. Only for the high-proficiency group of high school students, self-confidence was higher for long passages than for short passages (short = 3.303 and long = 3.375). For the other two groups of students (Group Middle and Low), self-confidence was not higher for the long passages in comparison to the short ones. Their self-confidence was higher regarding the short passages (Group Middle = 2.312 and Group Low = 1.724) whereas it was lower with the long passages (Group Middle = 2.075 and Group Low = 1.706). The difference of the high school students' self-confidence level was significantly higher for Group Middle for the shorter passages than the long ones.

4.2.2. Sensed Easiness

A paired samples *t*-test was conducted to investigate the effect of text length on the EFL readers' sensed easiness of the reading passages. The dependent variable was composed of scores of the students' answers on their sensed easiness of the reading passages. To ascertain the sense easiness of the readers while engaged with each passage, the readers were asked the phrase "I felt the reading passage was easy to understand" and their responses were recorded using a five-point Likert response scale. The results of the descriptive statistics of means and the standard deviation of the students' sensed easiness are shown in Table 4.8 below.

Table 4.9
Results for Sensed Easiness in Two Versions of the Passages by Proficiency Group

College students							N=100
Proficiency group	N	Text length				<i>p</i> value by <i>t</i> -test	
		Short		Long			
		M	SD	M	SD		
High	34	2.382	1.392	3.338	.935	.000***	
Middle	33	2.333	1.216	3.136	.945	.000***	
Low	33	2.06	1.073	2.984	.592	.000***	

Note. Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

High school students							N=102
Proficiency group	N	Text length				<i>p</i> value by <i>t</i> -test	
		Short		Long			
		M	SD	M	SD		
High	33	3.636	1.017	3.606	.826	.852	
Middle	40	2.625	1.107	2.312	1.029	.002**	
Low	29	1.982	.986	1.741	.751	.191	

Note. Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

As shown in Table 4.8, for the college students, the same trend of difference was also noted between the short and long text as the previous perception survey question (see 4.2.1). The college students' sensed easiness was higher for the long passages than for the short passages in all three groups (Group High: short = 2.382 and long = 3.338; Group Middle: short = 2.333 and long = 3.136; and Group Low: short = 2.06 and long = 2.984). In all three groups of the college students, the differences of the sensed easiness between the short and long passages were

statistically significant. On the other hand, for the high school students, their sense of easiness of the passages was quite different from the college students. In all three groups of the high school students, the sensed easiness seemed to be higher with the short passages whereas it was lower with the long passages (Group High: short = 3.636 and long = 3.606; Group Middle: short = 2.625 and long = 2.312; and Group Low: short = 1.982 and long = 1.741). However, the difference was statistically significant only for Group Middle.

4.2.3. Sufficiency of Information

To analyze the effect of text length on the readers' perception on the sufficiency of information given in the passage, a paired samples *t*-test was also employed. The sufficiency of information in each passage was asked with the survey question, "I felt that reading passage was long enough to understand what the author is trying to say" using a five-point Likert response scale. The results of the descriptive statistics of means and the standard deviation of the students' perception on the sufficiency of information are shown in Table 4.9.

Table 4.10
Results for Sufficiency of Information Two Versions of the Passages by Proficiency Group

College students							N=100
Proficiency group	N	Text length				<i>p</i> value by <i>t</i> -test	
		Short		Long			
		M	SD	M	SD		
High	34	3.235	.955	3.75	.939	.004**	
Middle	33	2.984	1.011	3.424	.969	.019**	
Low	33	2.666	.844	3.439	.726	.000***	

Note. Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

High school students							N=102
Proficiency group	N	Text length				<i>p</i> value by <i>t</i> -test	
		Short		Long			
		M	SD	M	SD		
High	33	3.681	.899	3.833	.932	.365	
Middle	40	2.687	1.054	2.825	1.163	.298	
Low	29	2.362	1.084	2.155	1.173	.322	

Note. Significance level: * $p < .05$, ** $p < .01$, *** $p < .001$

As shown in Table 4.9, in all three proficiency groups of college students, their perception of sufficiency of information was higher for the long passages than for the short passages (Group High: short = 3.235 and long = 3.75; Group Middle: short = 2.984 and long = 3.424; and Group Low: short = 2.666 and long = 3.439). There was a main effect of text length for all three groups, indicating that the college students perceived the information given in the text, to be sufficient

enough when the text was longer.

The same trend of difference was also noted for the high school students. For Group High and Group Middle, the sufficiency of information was higher for longer passages than for short passages (Group High: short = 3.681 and long = 3.833; Group Middle: short = 2.687 and long = 2.825), but for Group Low, the sufficiency of information was lower for longer passages (2.155), and higher for shorter passages (2.362).

4.3. Stimulated Recall Interview

To answer the third research question (“What are the reading strategies and generated inferences reported by the Korean EFL readers while reading a short passage and a long passage?”), the stimulated recall interview was analyzed. This section reports on a more detailed examination of the inference generation processes and strategy use employed by the EFL readers while they were trying to report their thought process with two different reading contexts (short and long passages).

For the analysis of the learners’ reading process, firstly, a multiple-choice question type (a gap-filling item) was used since it was considered to be appropriate to analyze the sensitiveness of the learners’ reading process to length of a given text. Secondly, open-ended questions (bridging inference and

elaborative inference) were used to examine the readers' ability to comprehend the overall arguments or main ideas which are implied, but not explicitly stated in the text. The analysis focused on the comparison of the two test formats: multiple-choice (a gap-filling item) and open-ended question (bridging and elaborative inference).

The stimulated recall interview was conducted two weeks after the main experiment. During the whole interview, the researcher met with 4 participants individually: 2 participants from each of the high and low group, all of whom were college students. Each individual appointment lasted approximately 30 to 40 minutes. Firstly, the participant was asked to read the short passage which they didn't get the correct answer to in the previous experiment and they were asked to report verbally what they were thinking immediately after they had read aloud each sentence. Then, they were asked to read the long version of same passage (one or two more paragraph was added to the short one) and encouraged to report any thoughts that came up to their minds and to explain any difference it had with the short version of the passage in detail. When any points made were unclear and needed clarification, they were interrupted by the researcher using the following questions:

- What do you mean by "... " (specific terms used in the passage)?
- Can you elaborate on the phrase, "... "?
- Why do you think the author used the expression, "... "?
- Which sentence do you think is the topic sentence and why?

The retrospective interview with the participant enabled the researcher to ascertain the actual reading comprehension processes which the learners are going through to find the answers. The meetings were audio-recorded and were later transcribed verbatim. Since the participants used Korean, their first language, the following interview protocol of the EFL learner were translated from Korean into English.

4.3.1. Reading Process for Multiple-Choice Questions

Participant sixty is an upper advanced level student who scored 928 on TEPS majoring in business administration. Her score on the reading comprehension measure was 13 on a scale of 16 points (81.25%). She scored 6 points for the short passages and 7 points for the long ones in total. For the multiple-choice questions, however, she scored high on the short passages, but scored lower on the longer passages. The researcher was curious to know the discrepancy between the participant's English proficiency level and her performance in the multiple-choice questions. Consequently, she was given the two versions of the passage which got the wrong answer (Passage no. 1) and was asked to verbalize her thoughts while reading the short passage.

The second interviewee, participant five is a lower-intermediate student (612 on

TEPS) majoring in architectural environment. Her score on the reading comprehension measure was 10 out of 16 points (62.5%). She got high scores (7 points) on the long passages, but low (3 points) on the short passages. She was also given the two versions of the same passage which got the wrong answer in the previous experiment (Passage no. 2) and reported her thinking process.

4.3.1.1. Short Passage

The major issue found in the multiple-choice question with the short version is that both of the readers were not sure of what should be filled in the blank regardless of their level of comprehension of the text as shown in Example 1 and 2. Both reported that the last two sentences of the passage should contain the main idea of the passage to find the best option, but the actual level of understanding of the two readers were quite different. The upper advanced level reader was very clear and exact about what she just read, whereas the lower intermediate student reported that she had no idea and had to look through the options to find the answer.

Example 1: Participant sixty

Even though I understand the overall meaning of the passage quite clearly, I am *still not sure* what would be the right option for the blank. So I mainly used the *elimination strategy* (Test-management) to find the most appropriate answer. Then, I reread the sentence after filling the gap with the option that I chose (Test-wiseness).

Example 2: Participant five

I am *really confused*. I read through the text (Reading strategy) and translate the sentences one by one (Reading strategy), but it is *still hard to understand* the overall meaning.

(Researcher: Can you tell me why you chose this option?)

I chose this option by eliminating the others (Test-wiseness). I think I could get the overall meaning of the text, but it was *a little bit confusing*, so I *couldn't be sure* about the answer I chose.

As shown in the examples above, since both of the readers found it difficult to choose the best one among five choices presented in the test, they had to resort to the test-taking strategies. This pattern of strategy use can be interpreted to mean that the multiple-choice questions do not stimulate the actual reading processes which the test-developers intended and test takers can guess the right answer without fully understand the reading passage, which is in direct line with the previous research studies (Nevo, 1989; Cohen, 1984; Anderson et al., 1991), as quoted above.

4.3.1.2. Long Passage

One of the issues found in the multiple-choice question with the long version is that in spite of the sufficiency of information given to the readers, ironically, both of the readers still were not sure of what could be the best option for the blank, as shown in the examples below.

Example 3: Participant sixty

What the author is really trying to say is *more clearly presented* in the long

version of the passage, but just because the idea of the author is presented *repeatedly* in the long text doesn't mean that the longer text hold more clues needed to fill the blank. I am *still confused* as to what should be in the text.

Example 4: Participant five

Before I read the longer text, I couldn't get the first sentence of the short passage "Some distinctions between good and bad are hardwired into our biology." *Now*, I really *know* what this sentence means. More information definitely helped me to better understand the main idea of the passage, but that made it *even more difficult* to find the answer to the multiple-choice question. Because there are more ideas, I am *more confused*.

As shown in Example 3, both of the readers made comments about the usefulness of abundant information in the longer text for reading comprehension. However, the upper advanced student felt that both of the short and long passages do not hold the clues necessary for the readers to fill in the blank. The lower-intermediate student also reported that more information in the longer text made her too confused to choose the best option to fill in the blank, as well. Their comments can be interpreted to be that the multiple choice question types, specifically gap-filling items do not necessarily relate to the length of the text. It seems that other factors, such as the students' background knowledge play more of an important role in finding the answers to these kinds of questions.

4.3.2. Reading Process for Open-Ended Questions

To examine whether the other participants' quality of reading process has been

changed by the length of the text given, two more interviewees were invited. Participant thirty-eight was the third interviewee, who is an upper advanced college student (930 on TEPS), majoring in geographic education. As with participant sixty mentioned above, she obtained higher scores on the short passages and lower scores on the long passages. She read the two versions of the passage in which she got the wrong answers previously (Passage no. 2) and was asked to verbalize her thoughts while reading the short passage.

The last interviewee, participant forty-four, is a lower-intermediate student (656 on TEPS) majoring in biology. She was given the literature passage in the short version and then the longer version. She was asked to report her thought process, accordingly.

As shown from Example 5 and 6, the trend observed in the multiple-choice questions occurred. Both of the readers reported that they benefited more from reading the longer version of the passages than the shorter ones when it comes to understanding the overall structure of the passage and natural flow of the ideas.

Example 5: Participant thirty-eight

The long version of the passage *helped me to understand* the overall structure of the reading passage. The two paragraphs consisted of two ideas contrasted and compared to each other. The additional paragraph in the long passage focused on the idea of a bad experience and the second paragraph is about the pressure, which *became clear* after reading the long version of the passage.

Example 6: Participant five

I think this passage (no. 4) is *much better* with the additional paragraph because reading the rest of the text *helped me to reach* a final understanding of

the passage based on *more clues presented* in the additional paragraph.

Even though the upper advanced level reader reported that she could not take advantage of the longer passage when it came to the multiple-choice items, the lower-intermediate reader did mention that the longer reading passage made it easier for her to comprehend the text and she was able to feel confident about her answers to the open-ended questions.

In sum, the participants' responses in stimulated recall interviews revealed some critical information to indicate that the more text can better assist the EFL readers' comprehension process. These individual points will be discussed in the following chapter.

CHAPTER 5.

DISCUSSION

In this chapter, the major findings obtained are presented with reference to each of the three research questions. The results of the current study are also discussed with relation to previous studies. The purpose of the thesis was to investigate the influence of text length and question type on Korean EFL learners' reading comprehension test performance. For this goal, Section 5.1 and 5.2 discusses the effects of text length and question type on the test-takers' performance and perceptions on the reading comprehension test, respectively. Section 5.3 presents the reading and test-taking strategies and inference generation that the Korean EFL learners employ while they are reading a short and long passage. In this section, the results from the statistical analyses are integrated with the stimulated recall interview data to explore the findings from the quantitative results in more detail through the qualitative data.

5.1. Effects of Text Length on Reading Comprehension

For the first research question ("Do the text length and the question type have an influence on Korean EFL readers' reading comprehension performance? If so, to what extent do the two variables have an effect on Korean EFL readers' reading

comprehension?”), the scores of the Korean EFL college and high school students' reading comprehension test by three different proficiency groups were examined.

Firstly, a paired samples *t*-test was conducted in order to test the main effect of the text length on the Korean EFL learners' overall reading comprehension test performance. The results showed that the readers' text comprehension is influenced by the length of the passages. The significant effect of the text length on readers' performance was found for both of the college and high school students. The statistical analysis indicated that the advanced and intermediate-level college students performed significantly better on the long version of the reading comprehension tests. Likewise, the advanced-level high school students benefited from the long passages more than the short ones. When only the scores of the multiple-choice questions were taken into account, the results of the present study seemed to be in agreement with the conclusion of the previous research studies (Hashemi & Bagheri, 2014; Lee, 1999), which claimed that text length did not have any significant effect on the test-takers' reading comprehension ability. However, the present study has shown different results to that of the previous research studies by measuring the students' reading comprehension using two different question types: multiple-choice and open-ended questions.

Although the results of the other proficiency groups did not meet the statistically meaningful level ($p < .05$), the overall trend of the results of the lower-level college students and all groups of the high school students clearly showed that the readers' performance on the longer passages were better than that of the short passages. The findings of the present study lend support to Cha's (1995) acknowledgement

of the effect of text length on the readers' comprehension performance. As Cha (1995) has pointed out, the reader's text comprehension is better when the texts became longer because the longer texts "provide more contextual cues, or information about the textual pattern" (p. 258). Even though the researcher did not utilize the open-ended questions, she made a clear distinction between the semantic and syntactic knowledge when it comes to making the reading comprehension assessment. In her study, both of the items were used to build experimental reading comprehension questions in an attempt to test the reader's ability of detecting deep structures embedded in the text. Similarly, in the present study, the open-ended inference generation questions were included, which ultimately promote a deeper thinking process in the reader, such as, being aware of the author's purpose and integrating prior knowledge with the text information. Therefore, by revealing the powerful effect of the text length on the readers' comprehension process, the present study underscores the importance of the length of the text used for the reading comprehension tests.

A repeated measure two-way ANOVA and an additional simple main effect analysis showed more important and interesting results with regard to the effects of text length and question type; the interaction effect between the two variables (text length and question type) was found. The interaction effect between the two variables was significant for the college students with a higher proficiency level of English: the advanced group (Group High) and the intermediate group (Group Middle). This indicates that text length and question type not only have significant effects on reading comprehension performance separately, but that they also

interact with each other as shown previously (see Section 4.1.1).

This interaction reflects the finding that, in the multiple-choice questions, the mean scores were marginally higher in the short texts than in the long texts, but the difference was not statistically significant. In other words, the reading comprehension performance measured by the multiple-choice questions did not make any difference between the short texts and the long texts. This suggests that giving more text did not help their reading comprehension performance measured by the multiple-choice questions. It can be inferred from the findings that, for the multiple-choice questions, which only require a selective response, a longer text may not be necessarily helpful for the readers to find the most likely option to fill in the blank. In other words, just because the readers read the additional text does not necessarily mean that they could choose the best option, which would fit into the blank. Rather, giving more text to the readers might have caused them to be distracted by the unnecessary text, which does not hold any useful information to arrive at a certain answer. In addition, as the gap-filling multiple-choice questions used in the present study (see Appendix A), the distractors of the multiple-choice questions are often constructed to be as attractive as possible. Therefore, the difficulty of the vocabulary in the answer options is often higher than the one in the passage, so the unfamiliar words used in the options might have caused the readers to experience difficulty in choosing the answer.

On the other hand, in the open-ended questions, the mean scores of the longer passages were significantly increased compared with the short ones. The results showed that the advanced-level and intermediate-level college students performed

significantly better with the long passages than the short ones. This is somewhat consistent with findings that when the students are given a longer text, their comprehension became greater. One might speculate that it might have been easier for the readers to construct their responses when they were reading longer texts with more redundant information presented in the text. This may also suggest that when reading comprehension is assessed through open-ended questions, the text that has a more coherent structure, which is a longer text in the present study, can enhance the readers' comprehension process.

Based on the finding presented above, it can be noted that the test-takers' performance on the multiple-choice questions did not really match with the results of the open-ended questions. The findings are in line with the previous study conducted by Ozuru, Brine, Kurby, and McNamara (2013), which concluded that open-ended and multiple-choice format questions measure different aspects of the comprehension processes. Ozuru and his colleagues have speculated that relying only on the multiple-choice question items to measure students' reading comprehension is not sufficient, but open-ended questions should be considered because they are "more sensitive than multiple-choice questions in measuring the quality of active generative processing of relevant and accurate ideas" (p.222). On this basis, the type of questions that induce more genuine understanding of the reading comprehension process to measure the readers' text comprehension should be used when designing reading comprehension assessments.

To recapitulate, the findings presented in the current study confirm that text length and question type did affect the reading comprehension performance of the

Korean EFL readers of different proficiency levels differently.

5.2. Effects of Text Length on Readers' Perception

For the second research question of Korean EFL readers' perception on the two different reading passages ("What is the Korean EFL readers' perception of the two different length of the passages?"), the readers' self-assessment survey was examined. The results showed the significant effect that the text length had on the college students' perception of the reading comprehension tests. This suggested that the college students perceived that their reading comprehension benefited from the long passages rather than that of the short ones in terms of their self-confidence, sensed easiness, and sufficiency of information.

However, the effect of text length on the perception of the high school students was found to be quite different from its effect on the college students. It seemed that the Group Middle and the Group Low of the high school students slightly preferred the short passages to the long passages. Even though there was no statistically significant difference between the students' perception of the two different reading contexts, there has been a trend showing that the high school readers, specifically the students in the low-proficiency group, appeared to feel less confident about their comprehension of the longer passages than the short ones.

According to the results of the Lexile measure (see Section 3.2), the short version of the passages is actually more difficult than the long version in three cases (passages no. 2, 3, and 4) and only in the case of the first passage did it lead to an increase in the readability compared to the longer passage (short passages: 1270L, 1190L, 820L, and 940L; long passages: 1090L, 1290L, 920L, and 970L). Based on the results of the text difficulties measured by Lexile, one would think that the students would perform better on the short version of the test than its longer version. However, the researcher predicted that even though the longer version had a greater number of ideas and longer sentences than the shorter version, the readers' performance would be better with longer passages with additional context, more redundancy and repeated information. As shown from the findings, the reading scores of both the college and high school students were better for the longer passages.

However, it was interesting to note that the high school EFL readers' perception of the reading passages did not exactly match with the results of the reading comprehension test scores. The high school students' self-assessment showed that they felt their reading comprehension was better when dealing with short passages than when they had to deal with the long ones, whereas their actual reading comprehension scores for the longer passages were better compared to the short passages.

One possible interpretation of the results is that, even though the readability indexes show that the difficulty of the reading passages are similar with that of their English textbooks, the high school students might not have been fluent

enough to read the materials as thoroughly as the researcher expected. Since the experimental passages of the present study were taken from the actual 2015 and 2016 Korean CSAT, the high school readers may have been too overwhelmed by the amount of unfamiliar words and lengthy texts to attend to the reading comprehension. Provided that they were high school freshmen and sophomores, the passages used for the present study might have been too difficult for the high school students. In other words, the high school students of lower proficiency may not have been able to demonstrate their English proficiency by completing the reading comprehension tests.

As Lee (2014) noted, it seems that little systematic attention has been paid to the discrepancy between text difficulty measure and the readers' actual reading fluency within the Korean EFL setting. Although a certain level of difficulty that the readers are supposed to read when they get to a certain grade level is presumably established by the educational authorities (often presented through English textbooks), there is little empirical data to support the notion that students' actual reading performance increases as their grade level rises.

The findings gathered from the high school students build strong support for the incorrectness of readability indexes. The results are in line with Yi (2013) who claimed that passage length is a crucial factor for the test developers to take into account and thus, additional procedures for checking the level of difficulty of the text should be considered to guarantee the comprehensibility of the reading passages, such as in-depth qualitative examination of text comprehensibility.

5.3. Readers' Strategy Use on Reading Performance

From the stimulated recall interviews regarding the Korean learners' reading comprehension, the third research question ("what are the reading strategies and generated inferences reported by the Korean EFL readers while reading a short passage and a long passage?") was examined. One of the issues found is that the multiple-choice questions (gap-filling items) of both versions of the passages did not stimulate the actual reading process in a similar way to that of the authentic academic reading process. Both groups of interviewees (high and low-intermediate proficiency readers) reported that the given passage did not have any direct clues to help them choose the best option out of five choices. All of those distractors seemed to have key words in the text, which made them even more confused. This was especially the case with the lower-intermediate student. She said that she was forced to choose the answer somehow by relying on test-wisness strategies.

The findings replicate the results of Nam's (2015) qualitative research. She explored Korean EFL readers' test-taking strategy use in gap-filling items for the CSAT through analyzing think-aloud data. Her study had similar results, in that test-takers quickly turned to test-wisness and test management strategies using only fragmented parts of the text when they failed to understand the main idea of the reading passages. This is also consistent with the results of Cordon and Day's (1996) study, which also showed that students in standardized test conditions

made significantly greater use of strategies than students in the main idea condition.

In addition, another issue found from the stimulated recall is that readers' level of comprehension of the reading passage improved after reading the passage again with the additional paragraph, especially, for the open-ended questions. Although the upper advanced reader said that she did not take advantage of the longer passages, the lower intermediate readers felt that reading the longer passage was helpful to understand the overall context of the text, which eventually led them to proper comprehension of the text. Based on the results, it can be speculated that giving more text may enhance the EFL readers' comprehension process.

CHAPTER 6.

CONCLUSION

This chapter is composed of three sections. Section 6.1 summarizes the major findings of the present study. In Section 6.2, the pedagogical implications for reading assessments are presented. Finally, Section 6.3 reports the limitations of the current study and makes suggestions for future research.

6.1. Major Findings

The present study was designed to explore the effect that the length of the passages has on EFL readers' performance during standardized tests and their perception of the two different text types. This study dealt with three key points in the research questions: 1) Do the text length and the question type have an influence on Korean EFL readers' reading comprehension performance? If so, to what extent do the two variables have an effect on Korean EFL readers' reading comprehension? 2) What is the Korean EFL readers' perception of the two different lengths of the passages? 3) What are the reading strategies and generated inferences reported by the Korean EFL readers while reading a short passage and a long passage? The key findings of the current study are summarized below.

Regarding the first research question of whether text length and question type

have an influence on Korean EFL readers' performance and perception, there was a significant difference between the two types of the tests. A repeated measure two-way ANOVA revealed that for the college students, the advanced and intermediate-level students performed significantly better with the longer version of the reading comprehension tests. And it was also suggested that the question types variable affected all three proficiency groups of college students, meaning that all three of the college EFL readers obtained higher scores for the open-ended questions rather than the multiple-choice questions.

A similar trend occurs among high school students. All of the proficiency groups of the high school students performed marginally better on the long version of the reading tests, but the results of a repeated measure two-way ANOVA showed that the mean difference between the short and long version was statistically significant for the high proficiency group of high school students.

For the second research question of whether the text length has an impact on the readers' perception on the two tests, there was a main effect of text length for the college students on their perception of the reading comprehension tests in terms of their self-confidence, sensed easiness, and sufficiency of information. Findings confirmed that the college students perceived that their reading comprehension benefited from the long passages more than that of the short ones. However, for the high school students, they seemed to prefer the short passages to the long ones. The difference, however, was marginal and did not have any statistical significance.

Lastly, concerning the third research question of what the reading strategies and

generated inferences are reported by the Korean EFL readers while reading a short passage and a long passage, the findings confirmed that the multiple-choice questions (gap-filling items) did not stimulate the actual reading process and that it actually leads less-competent readers to rely on test-management and test-wiseness strategies. The findings also suggest that multiple-choice and open-ended questions assess different reading processes and each of them tap into distinctive processes. It seems that open-ended questions reflect the readers' true reading abilities.

6.2. Pedagogical Implications

6.2.1. Finding the Right Reading Texts

The reading passages used for the standardized reading comprehension tests administered in Korean have been extracted from truncated parts of research papers, novels, and even professional literature. The selection of the passages is mainly based on their linguistic difficulties (syntactic complexity or vocabulary difficulty) or test-developers' personal preference. As shown from the stimulated recall interview above, the readers reported that they could better understand the overall structure of the long passages and that the unnatural structure of short passages inhibit the readers' processes of reading comprehension. Consequently,

the readers struggle to understand the meaning of the texts by searching of the text for mission information, or using their background knowledge to integrate the overall meaning of the text. If they are in the standardized test situation, where they are forced to choose one and only option, which eventually lead the readers to rely on test-taking strategies, and not on authentic academic reading process.

Unfortunately, there have been little attempts to investigate the effect of using truncated passages on EFL learners' reading comprehension processes so far. However, the findings of the present study clearly suggest that it is essential to know what the impact of the amount of text used for reading during comprehension tests might be. As with any step away from conventional practice, there must be some potential difficulties, but it is crucial that test-developers take these findings into account because the impact is particularly significant when it comes to learners of lower language proficiency.

6.2.2. Selection of Question Types for Reading Tests

The results of the present study suggest that two different reading question types (multiple-choice and open-ended) induce and tap into distinct comprehension processes. The findings are of particular importance to the development of reading comprehension tests. As Bachman noted (1990), in order to interpret the performance of assessment as an indicator of the test-taker's reading abilities

which we want to measure, we should first clearly identify and define the construct of the reading comprehension abilities. That is, if the test developers' expectations from the respondents do not match with the actual processes that test-takers undergo during testing, we cannot say that the test is actually measuring the reading comprehension ability that we intended to measure. This eventually affects the validity of inferences we make on the basis of the test scores. Especially, if the results of the reading scores have a serious consequence on the stakeholders, all questions types included in the reading assessment tasks need to go through a careful validation process, such as an in-depth qualitative analysis of questions so that they could interpret the test-takers' scores with regards to the English reading comprehension section of the CSAT as their real academic reading abilities.

6.3. Limitations and Suggestions for Future Research

The results of the study have illustrated the effects of text length and questions types of the reading comprehension tests on Korean EFL readers' performance, perception and their strategy use. However, the present study has some limitations that need to be considered regarding the methodology despite the fact that each procedure in the experiment and the stimulated recall interview were carefully designed.

Firstly, with the concern of the cognitive load that the participants might

experience, only four passages have been used for reading comprehension tests. Such a decision was based on the mini-pilot study prior to the main experiment to assess the viability of the reading comprehension tests. This included text readability, the number of questions, and the adequacy of test level for the Korean EFL students. Since the present study is limited by the small number of test items of each type, an additional experiment for the high school students was conducted in an effort to increase the meaningful results of the multiple-choice reading comprehension tests. Nevertheless, different total scores for each question type would make it arguable to compare the results of multiple-choice and open-ended questions directly. Therefore, in future research, it would be suggested to examine the effects of text length and questions types on EFL readers' performance by balancing the number of each question type.

Next, only four college students participated in the stimulated recall interview in order to investigate the EFL readers' thinking process while they were engaged in the two different reading contexts. Although the interviews revealed some insightful issues regarding the amount of the reading text and item types for reading comprehension tests, more participants would have produced more convincing evidence to support the claims made in the study.

Lastly, the participants' reading proficiency levels were based on their standardized reading comprehension test scores obtained previously and the self-reported proficiency level was not a perfect match for the performance of the reading comprehension tests for the present study. Therefore, this mismatch between the two contexts could be resolved in future research by conducting

specialized reading comprehension tests and comparing the groups with pretest and posttest data.

REFERENCES

- 이병민. (2014). *당신의 영어는 왜 실패하는가?: 대한민국에서 영어를 배운다는 것*. 서울: 우리학교.
- Akamatsu, N. (2002). A similarity in word-recognition procedures among second language readers with different first language backgrounds. *Applied Psycholinguistics*, 23(1), 117-133.
- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *The Modern Language Journal*, 75(4), 460-472.
- Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Beach, K. L. (2008). *The effect of media, text length, and reading rates on college student reading comprehension levels*. (Unpublished doctoral dissertation). University of Phoenix, Tempe, AZ.
- Bernhardt, E. B. (1991). *Reading development in a second language: theoretical, empirical, and classroom perspectives*. Norwood, NJ: Ablex Publishing Corporation.
- Campbell, J. R. (1999). *Cognitive processes elicited by multiple-choice and constructed-response questions on an assessment of reading comprehension*. Unpublished doctoral dissertation, Temple University, Philadelphia, PA, USA
- Cha, K. A. (1995). The effect of text length in the diagnosis of reading comprehension. *Journal of the Applied Linguistics Association of Korea*, 8(1), 249-275.
- Chastain, K. (1988). *Developing Second Language Skills: Theory and Practice*. Orlando, Florida: Harcourt Brace Jovanovich.
- Choi, M. S. (2011). *Analysis of the factors affecting passage dependency of multiple-choice English reading comprehension tests*. (Unpublished master's thesis). Ewha Women's University, Seoul, Korea.
- Cohen, A. D. (1984). On taking language tests. *Language Testing*, 1(1), 1.
- Cohen, A. D. (2014). *Strategies in learning and using a second language*. New York: Routledge.
- Cohen, A. D., & Upton, T. A. (2007). 'I want to go back to the text': Response strategies on the reading subtest of the new TOEFL®. *Language Testing*, 24 (2), 209-250
- Commander, N. E., & Stanwyck, D. J. (1997). Illusion of knowing in adult

- readers: Effects of reading skill and passage length. *Contemporary Educational Psychology*, 22(1), 39-52.
- Cordón, L. A., & Day, J. D. (1996). Strategy use on standardized reading comprehension tests. *Journal of educational psychology*, 88(2), 288-295.
- Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse processes*, 25(1), 1-53.
- Crossley, S. A., & McNamara, D. S. (2008). Assessing L2 reading texts at the intermediate level: An approximate replication of Crossley, Louwse, McCarthy & McNamara (2007). *Language Teaching*, 41(3), 409-429.
- Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General*, 130(2), 208-223.
- Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first-and second-language learners. *Reading Research Quarterly*, 38(1), 78-103.
- Flippo, R. F., & Schumm, J. S. (2000). Reading tests. In R. F. Flippo & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research* (pp. 403-472). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Fukkink, R. G., Hulstijn, J., & Simis, A. (2005). Does training in second-language word recognition skills affect reading comprehension? An experimental study. *The Modern Language Journal*, 89(1), 54-75.
- Goodman, K. S., & Freeman, D. (1993). What's simple in simplified language? In M. L. Tickoo (Ed.), *Simplification: Theory and application* (pp. 69-76). Singapore: SEAMEO Regional Language Center.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.
- Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 3-26). New York: Lawrence Erlbaum Associates.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing Inferences During Narrative Text Comprehension. *Psychological Review*, 101(3), 371-395.
- Hashemi, S. Z., & Bagheri, M. S. (2014). External factors influences on EFL learners' reading comprehension test performance. *International Journal of Language Learning and Applied Linguistics World*, 7(1), 150-166.
- Hogan, T., Bridges, M. S., Justice, L. M., & Cain, K. (2011). Increasing higher level language skills to improve reading comprehension. *Focus on Exceptional Children*, 44(3), 1-20.
- Hughes, A. (2003). *Testing for language teachers*. New York: Cambridge

University Press.

- Jalilehvand, M. (2012). The effects of text length and picture on reading comprehension of Iranian EFL students. *Asian Social Science*, 8(3), 329.
- Johnston, P. (1984). Prior knowledge and reading comprehension test bias. *Reading Research Quarterly*, 19(2), 219-239.
- Jones, L. (2005). *The effect of applying active reading study strategies to varying text lengths on the lower-level comprehension of developmental reading students* (Unpublished doctoral dissertation). University of Houston. USA.
- Katz, S., Lautenschlager, G. J., Blackburn, A. B., & Harris, F. H. (1990). Answering reading comprehension items without passages on the SAT. *Psychological Science*, 1, 122-127.
- Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. (2007). Comprehension in preschool and early elementary children: Skill development and strategy interventions. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 27-45). New York: Lawrence Erlbaum Associates.
- Kim, S. H., & Chon, Y. V. (2014). Test-taking strategies of L2 adolescent learners: Three multiple-choice items and L2 proficiency. *English Teaching*, 69(1), 61-90.
- Kim, S. J. (2016). *Analyzing the cohesion of the reading passages on standardized English tests: CSAT, TEPS, EIKEN and TOEFL* (Unpublished master's thesis). Seoul National University, Seoul, Korea.
- Kim, Y. C. (2003). The effects of strategic processing in reading upon inference generation and comprehension. *The International Journal of Creativity & Problem Solving*, 13(1), 31-48.
- King, A. (2007). Beyond literal comprehension: A strategy to promote deep understanding of text. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 267-290). New York: Lawrence Erlbaum Associates.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review*, 95(2), 163.
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49(4), 294.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge university press.
- Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209-226). Malden, MA: Blackwell Publishing.
- Kispal, A. (2008). *Effective teaching of inference skills for reading: Literature review. Research Report DCSF-RR031*. London: DCSF.
- Koda, K. (1998). The role of phonemic awareness in second language reading. *Second Language Research*, 14(2), 194-215.

- Koda, K. (1999). Development of L2 intraword orthographic sensitivity and decoding skills. *The Modern Language Journal*, 83(1), 51-64.
- Kobayashi, M. (2002). Method effects on reading comprehension test performance: Text organization and response format. *Language Testing*, 19(2), 193-220.
- Lee, K. S. (1999). The effects of the number of questions per passage, the length of passage, and the topic familiarity on multiple-choice English listening and reading comprehension tests. *English Teaching*, 54(4), 327-351.
- Lee, J. W., & Ku, M. (2005). A case study on the use of English reading test-taking strategies. *English Language Teaching*, 17(2), 179-201.
- Magliano, J. P., Millis, K., Ozuru, Y., & McNamara, D. S. (2007). A multidimensional framework to evaluate reading assessment tools. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 107-136). New York: Lawrence Erlbaum Associates.
- Magliano, J. P., Millis, K. K., Levinstein, I., & Boonthum, C. (2011). Assessing comprehension during reading with the reading strategy assessment tool (RSAT). *Metacognition and Learning*, 6(2), 131-154.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99(3), 440-466.
- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14(1), 1-43.
- McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse processes*, 22(3), 247-288.
- Mehrpour, S., & Riazi, A. (2004). The impact of text length on EFL students' reading comprehension. *Asian EFL Journal*, 6(3), 1-13.
- Millis, K., Magliano, J., & Todaro, S. (2006). Measuring discourse-level processes with verbal protocols and latent semantic analysis. *Scientific Studies of Reading*, 10(3), 225-240.
- Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental psychology*, 40(5), 665.
- Nam, Y. K. (2015). Korean EFL students' strategy use in gap-filling inference items. *English Teaching*, 70, 81-107.
- Nevo, N. (1989). Test-taking strategies on a multiple-choice test of reading comprehension. *Language Testing*, 6(2), 199-215.
- Oakhill, J., & Cain, K. (2007). Issues of causality in children's reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 47-72). New

- York: Lawrence Erlbaum Associates.
- Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16*(2), 91-121.
- O'Reilly, T., & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "High-Stakes" measures of high school students' science achievement. *American Educational Research Journal, 44*(1), 161-196.
- Ozuru, Y., Rowe, M., O'Reilly, T., & McNamara, D. (2008). Where's the difficulty in standardized reading tests: The passage or the question? *Behavior Research Methods, 40*(4), 1001-1015.
- Ozuru, Y., Briner, S., Kurby, C. A., & McNamara, D. S. (2013). Comparing comprehension measured by multiple-choice and open-ended questions. *Canadian Journal of Experimental Psychology, 67*(3), 215.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. *Cognition and Instruction, 1*(2), 117-175.
- Park, C. (2014, Feb). *The college scholastic ability test in Korea: Introduction and related issues*. Paper presented at the KAERA research forum, Seoul, Korea.
- Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University Press.
- Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. S. Gorfein, (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67-86). Washington, DC, US: American Psychological Association
- Rasinski, T. V. (2012). Why reading fluency should be hot. *Reading Teacher, 65*(8), 516-522.
- Rothkopf, E. Z., & Billington, M. J. (1983). Passage length and recall with test size held constant: Effects of modality, pacing, and learning set. *Journal of verbal learning and verbal behavior, 22*(6), 667-681.
- Snow, C. E. (2002). *Reading for understanding: Toward a research and development program in reading comprehension*. Santa Monica, CA: Rand.
- Song, H. (1998). The effect of the college scholastic ability test (CSAT) on Korean college students' reading processes and their perceptions of reading in English. *English Teaching, 53*(4), 265-290.
- Stolarova, M., Wolf, C., Rinker, T., & Briemann, A. (2014). How to assess and compare inter-rater reliability, agreement and correlation of ratings: an exemplary analysis of mother-father and parent-teacher expressive vocabulary rating pairs. *Frontiers in Psychology, 5*, 509.
- Sternberg, R. J. (1991). Are we reading too much into reading comprehension tests? *Journal of Reading, 34*(7), 540-545.
- Taylor, B. M., & Samuels, S. J. (1983). Children's use of text structure in the

- recall of expository material. *American Educational Research Journal*, 20(4), 517-528.
- Todd, W. B., & Kessler, C. C. (1971). Influence of response mode, sex, reading ability, and level of difficulty on four measures of recall of meaningful written material. *Journal of Educational Psychology*, 62(3), 229.
- Urquhart, A. H., & Weir, C. J. (2014). *Reading in a second language: Process, product and practice*. New York: Routledge.
- Valencia, S. Hiebert, E. H., and & Afflerbach, P. (1994). Definitions and perspectives. In S. Valencia, E. H. Hiebert, & P. Afflerback (Eds.), *Authentic reading assessment: Practices and possibilities* (pp. 6-21). Newark, DE: International Reading Association.
- Valencia, S. W., & Pearson, P. D. (1988). Principles for classroom comprehension assessment. *Remedial and Special Education*, 9(1), 26-35.
- Van Den Broek, P., Kendeou, P., Kremer, K., Lynch, J., Butler, J., White, M. J., & Lorch, E. P. (2005). Assessment of comprehension abilities in young children. *Children's reading comprehension and assessment*, 107-130.
- Van Dijk, T. A., Kintsch, W., & Dijk, T. A. V. (1983). *Strategies of discourse comprehension*: New York: Academic Press.
- Yi, Y. S. (2013). On the Optimal Text Length of Reading Comprehension Tests. *Journal of the English Language and Literature*, 55(4), 505-530.
- Yoshida, M. (2003). Working memory capacity and use of inference in L2. *The Japan Association of College English Teachers (JACET) Bulletin*, 36, 1-17.
- Zwaan, R. A., & Singer, M. (2003). Text comprehension. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

APPENDICES

APPENDIX 1.	106
APPENDIX 2.	115

APPENDIX 1.

Test Form A

1.

The concept of humans doing multiple things at a time has been studied by psychologists since the 1920s, but the term “multitasking” didn’t exist until the 1960s. It was used to describe computers, not people. Back then, ten megahertz was so fast that a new word was needed to describe a computer’s ability to quickly perform many tasks. In retrospect, they probably made a poor choice, for the expression “multitasking” is inherently deceptive. Multitasking is about multiple tasks alternately sharing one resource (the CPU), but in time the context was flipped and it became interpreted to mean multiple tasks being done simultaneously by one resource (a person). It was a clever turn of phrase that’s misleading, for even computers can process only one piece of code at a time. When Ⓐthey “multitask,” they switch back and forth, alternating their attention until both tasks are done. The speed with which computers tackle multiple tasks that everything happens at the same time, so Ⓑcomparing computers to humans can be confusing.

1-1. 빈칸에 들어갈 말로 가장 적절한 것을 고르세요.

- ① expels the myth
- ② feeds the illusion
- ③ conceals the fact
- ④ proves the hypothesis
- ⑤ blurs the conviction

1-2. 밑줄 친 ㉠they 가 가리키는 것을 설명하세요.

.....

1-3. 이 글의 작가가 ㉡“comparing computers to humans can be confusing” 와 같이 주장하는 이유를 설명하세요.

.....

1-4. 위 글을 50자 내외의 우리말로 요약하세요.

.....

▣ 자기평가 설문

(1번 문항에 대한 설문에 답해주세요.)

Survey Questions	전혀 아니다 ←→ 매우 그렇다
1. 나는 지문의 내용에 친숙하다. (지문을 미리 읽어 보거나 공부한 적이 있음)	1 2 3 4 5
2. 나는 내가 선택한 답이 정답이라고 확신한다.	1 2 3 4 5
3. 위의 읽기 지문은 이해하기 쉽다.	1 2 3 4 5
4. 지문에 주어진 정보가 ‘빈 칸 채우기 문제’를 푸는데 충분하다고 생각한다.	1 2 3 4 5

2.

The psychologist Paul Rozin, an expert on disgust, observed that a single cockroach will completely wreck the appeal of a bowl of cherries, but a cherry will do nothing at all for a bowl of cockroaches. As he points out, the negative trumps the positive in many ways, and loss aversion is one of many manifestations of a broad negativity dominance. Other scholars, in a paper titled, “Bad Is Stronger Than Good,” summarized the evidence as follows: “Bad emotions, bad parents, and bad feedback have more impact than good ones, and bad information is processed more thoroughly than good. The self is more motivated to avoid bad self-definitions than to pursue good ones. Bad impressions and bad stereotypes are quicker to form and more resistant to disconfirmation than good ones.” They cite John Gottman, the well-known expert in marital relations, who observed that the long-term success of a relationship depends far more on avoiding the negative than on seeking the positive. Gottman estimated that a stable relationship requires that good interactions outnumber bad interactions by at least 5 to 1. Other asymmetries in the social domain are even more striking. We all know that a friendship that may take years to develop can be ruined by a single action.

Some distinctions between good and bad are hardwired into our biology. Infants enter the world ready to respond to pain as bad and to sweet (up to a point) as good. In many situations, however, the boundary between good and bad is a reference point that changes over time and depends on the immediate circumstances. Imagine that you are out in the country on a cold night, inadequately dressed for the pouring rain, your clothes soaked. A stinging cold wind completes your misery. As you wander around, you find a large rock

that provides some shelter from the fury of the elements. The biologist Michel Cabanac would call the experience of ㉠that moment intensely pleasurable because it functions, as pleasure normally does, to indicate the direction of _____. ㉡The pleasant relief will not last very long, of course, and you will soon be shivering behind the rock again, driven by your renewed suffering to seek better shelter.

* shiver: 떨다

2-1. 빈칸에 들어갈 말로 가장 적절한 것을 고르세요.

- ① a permanent emotional adjustment to circumstantial demands
- ② enhancing self-consciousness through physical suffering
- ③ a biologically significant improvement of circumstances
- ④ judging desirable and undesirable conditions impartially
- ⑤ a mentally pre-determined inclination for emotional stability

2-2. 밑줄 친 ㉠that moment 가 가리키는 것을 설명하세요.

.....

2-3. 이 글의 작가가 ㉡와 같이 ㉢“the pleasant relief will not last very long” 주장하는 이유를 설명하세요.

.....

2-4. 위 글을 50자 내외의 우리말로 요약하세요.

.....

■ 자기평가 설문

(1번 문항에 대한 설문에 답해주세요.)

Survey Questions	전혀 아니다 ←→ 매우 그렇다				
1. 나는 지문의 내용에 친숙하다. (지문을 미리 읽어 보거나 공부한 적이 있음)	1	2	3	4	5
2. 나는 내가 선택한 답이 정답이라고 확신한다.	1	2	3	4	5
3. 위의 읽기 지문은 이해하기 쉽다.	1	2	3	4	5
4. 지문에 주어진 정보가 ‘빈 칸 채우기 문제’를 푸는데 충분하다고 생각한다.	1	2	3	4	5

3.

My friend was disappointed that scientific progress has not cured the world’s ills by abolishing wars and starvation; that gross human inequality is still widespread; that happiness is not universal. My friend made a common mistake — a basic misunderstanding in the nature of knowledge. Knowledge is amoral — not immoral but morality neutral. It can be used for any purpose, but many people assume it will be used to further their favorite hopes for society — and ①this is the fundamental flaw. Knowledge of the world is one thing; its uses create a separate issue. To be disappointed that our progress in understanding has not remedied the social ills of the world is a legitimate view, but _____ . ② To argue that knowledge is not progressing because of the African or Middle Eastern conflicts misses

the point. There is nothing inherent in knowledge that dictates any specific social or moral application.

3-1. 빈칸에 들어갈 말로 가장 적절한 것을 고르세요.

- ① to confuse this with the progress of knowledge is absurd
- ② to know the nature of knowledge is to practice its moral value
- ③ to remove social inequality is the inherent purpose of knowledge
- ④ to accumulate knowledge is to enhance its social application
- ⑤ to make science progress is to make it cure social ills

3-2. 밑줄 친 ㉠this 가 가리키는 것은?

.....

3-3. 이 글의 작가가 ㉡“To argue that knowledge is not progressing ... misses the point” 와 같이 주장하는 이유를 설명하세요.

.....

3-4. 위 글을 50자 내외의 우리말로 요약하세요.

.....

▣ 자기평가 설문

(1번 문항에 대한 설문에 답해주세요.)

Survey Questions	전혀 아니다 ←→ 매우 그렇다				
1. 나는 지문의 내용에 친숙하다. (지문을 미리 읽어 보거나 공부한 적이 있음)	1	2	3	4	5
2. 나는 내가 선택한 답이 정답이라고 확신한다.	1	2	3	4	5
3. 위의 읽기 지문은 이해하기 쉽다.	1	2	3	4	5
4. 지문에 주어진 정보가 ‘빈 칸 채우기 문제’를 푸는데 충분하다고 생각한다.	1	2	3	4	5

4.

Long before Walt Whitman wrote *Leaves of Grass*, poets had addressed themselves to fame. Horace, Petrarch, Shakespeare, Milton, and Keats all hoped that poetic greatness would grant them a kind of earthly immortality. Whitman held a similar faith that for centuries the world would value his poems. But to this ancient desire to live forever on the page, he added Ⓐ a new sense of fame. Readers would not simply attend to the poet’s work; they would be attracted to the greatness of his personality. They would see in his poems a vibrant cultural performance, an individual springing from the book with tremendous charisma and appeal. Out of the political rallies and electoral parades that marked Jacksonian America, Whitman defined poetic fame in relation to the crowd. Other poets might look for their inspiration from the goddess of poetry. Whitman’s poet sought_____ . In the instability of American democracy, fame would be dependent on celebrity, on the degree to which Ⓑ the people rejoiced in the poet and his work.

This book tells the story of how an obscure Brooklyn poet, better known for his political journalism than verse, immersed himself in the culture of celebrity that was then emerging in the United States. Hoping to redress the mounting divisions in his country, he declared that the poet would become the center of American civic life, that he would command more power and sway than the political representatives he expected to supersede. As Whitman imagined it, the story of celebrity would be the story of democracy. He hoped that the nation's narrow political institutions would undergo an extraordinary transformation once they encountered the populist power embodied in the poet's fame.

Whitman would have to wait a long time for his celebrity to appear, and even then, it came not in the overflow of spontaneous affection he had envisioned as a younger man but in the respect and admiration expressed by his fellow dignitaries. Uncomfortable with the magnetic, sexual rebel depicted in the first editions of *Leaves of Grass*, the culture finally accepted Whitman as a majestic, grandfatherly poet, a patriotic wound dresser with this long white beard. But this book is not a study of Whitman's reputation, of how he was eventually accepted into the canon and became the complex national icon that he is today. I am interested in how Whitman's ideas of celebrity helped shape *Leaves of Grass*, how the meanings he granted to fame influenced the relationship he established with his readers. Sketched out in his notebooks, advertisements, conversations, and poetry, Whitman's celebrity was as much of a form of advance publicity as it was a biographical achievement.

4-1. 빈칸에 들어갈 말로 가장 적절한 것을 고르세요.

- ① poetic purity out of political chaos
- ② fame with political celebrities

- ③ a refuge from public attention
- ④ immortality in literature itself
- ⑤ the approval of his contemporaries

4-2. 밑줄 친 ㉠a new sense of fame 을 설명하세요.

.....

4-3. Whitman과 그의 시에 대해서 왜 대중이 ㉡와 같이 반응했는지 설명하세요.

.....

4-4. 위 글을 50자 내외의 우리말로 요약하세요.

.....

▣ 자기평가 설문

(1번 문항에 대한 설문에 답해주세요.)

Survey Questions	전혀 아니다 ←→ 매우 그렇다				
1. 나는 지문의 내용에 친숙하다. (지문을 미리 읽어 보거나 공부한 적이 있음)	1	2	3	4	5
2. 나는 내가 선택한 답이 정답이라고 확신한다.	1	2	3	4	5
3. 위의 읽기 지문은 이해하기 쉽다.	1	2	3	4	5
4. 지문에 주어진 정보가 ‘빈 칸 채우기 문제’를 푸는데 충분하다고 생각한다.	1	2	3	4	5

APPENDIX 2.

Test Form B

1.

Multitasking is a lie. It's a lie because nearly everyone accepts it as an effective thing to do. It's become so mainstream that people actually think it's something they should do, and do as often as possible. We not only hear talk about doing it, we even hear talk about getting better at it. More than six million webpages offer answers on how to do it, and career websites list "multitasking" as a skill for employers to target and for prospective hires to list as a strength. Some have gone so far as to be proud of their supposed skill and have adopted it as a way of life. But it's actually a "way of lie," for the truth is multitasking is neither efficient nor effective. In the world of results, it will fail you every time. When you try to do two things at once, you either can't or won't do either well. If you think multitasking is an effective way to get more done, you've got it backward. It's an effective way to get less done. As Steve Uzzellsaid, "Multitasking is merely the opportunity to screw up more than one thing at a time."

The concept of humans doing multiple things at a time has been studied by psychologists since the 1920s, but the term "multitasking" didn't exist until the 1960s. It was used to describe computers, not people. Back then, ten megahertz was so fast that a new word was needed to describe a computer's ability to quickly perform many tasks. In retrospect, they probably made a poor choice, for the expression "multitasking" is inherently deceptive. Multitasking is about multiple tasks alternately sharing one resource (the CPU), but in time the context was flipped and it became interpreted to mean multiple

tasks being done simultaneously by one resource (a person). It was a clever turn of phrase that's misleading, for even computers can process only one piece of code at a time. When they "multitask," they switch back and forth, alternating their attention until both tasks are done. The speed with which computers tackle multiple tasks that everything happens at the same time, so comparing computers to humans can be confusing.

1-1. 빈칸에 들어갈 말로 가장 적절한 것을 고르세요.

- ① expels the myth
- ② feeds the illusion
- ③ conceals the fact
- ④ proves the hypothesis
- ⑤ blurs the conviction

1-2. 밑줄 친 they 가 가리키는 것을 설명하세요.

.....

1-3. 이 글의 작가가 “comparing computers to humans can be confusing” 와 같이 주장하는 이유를 설명하세요.

.....

1-4. 위 글을 50자 내외의 우리말로 요약하세요.

.....

■ 자기평가 설문

(1번 문항에 대한 설문에 답해주세요.)

Survey Questions	전혀 아니다 ←→ 매우 그렇다				
1. 나는 지문의 내용에 친숙하다. (지문을 미리 읽어 보거나 공부한 적이 있음)	1	2	3	4	5
2. 나는 내가 선택한 답이 정답이라고 확신한다.	1	2	3	4	5
3. 위의 읽기 지문은 이해하기 쉽다.	1	2	3	4	5
4. 지문에 주어진 정보가 '빈 칸 채우기 문제'를 푸는데 충분하다고 생각한다.	1	2	3	4	5

2.

Some distinctions between good and bad are hardwired into our biology. Infants enter the world ready to respond to pain as bad and to sweet (up to a point) as good. In many situations, however, the boundary between good and bad is a reference point that changes over time and depends on the immediate circumstances. Imagine that you are out in the country on a cold night, inadequately dressed for the pouring rain, your clothes soaked. A stinging cold wind completes your misery. As you wander around, you find a large rock that provides some shelter from the fury of the elements. The

biologist Michel Cabanac would call the experience of ㉠that moment intensely pleasurable because it functions, as pleasure normally does, to indicate the direction of _____ . ㉡The pleasant relief will not last very long, of course, and you will soon be shivering behind the rock again, driven by your renewed suffering to seek better shelter.

* shiver: 떨다

2-1. 빈칸에 들어갈 말로 가장 적절한 것을 고르세요.

- ① a permanent emotional adjustment to circumstantial demands
- ② enhancing self-consciousness through physical suffering
- ③ a biologically significant improvement of circumstances
- ④ judging desirable and undesirable conditions impartially
- ⑤ a mentally pre-determined inclination for emotional stability

2-2. 밑줄 친 ㉠that moment 가 가리키는 것을 설명하세요.

.....

2-3. 이 글의 작가가 ㉡와 같이 ㉢“the pleasant relief will not last very long” 주장하는 이유를 설명하세요.

.....

2-4. 위 글을 50자 내외의 우리말로 요약하세요.

.....

▣ 자기평가 설문

(1번 문항에 대한 설문에 답해주세요.)

Survey Questions	전혀 아니다 ←→ 매우 그렇다				
1. 나는 지문의 내용에 친숙하다. (지문을 미리 읽어 보거나 공부한 적이 있음)	1	2	3	4	5
2. 나는 내가 선택한 답이 정답이라고 확신한다.	1	2	3	4	5
3. 위의 읽기 지문은 이해하기 쉽다.	1	2	3	4	5
4. 지문에 주어진 정보가 ‘빈 칸 채우기 문제’를 푸는데 충분하다고 생각한다.	1	2	3	4	5

3.

The conversation was lively, and we talked of recent medical discoveries exciting to me. Then my friend started matter-of-factly that he was “disappointed in science!” That caused me to ponder. I could not simply dismiss what he said. He is too thoughtful and sincere for that. Surely this modern, highly educated man did not mean he was disappointed in our discovering the composition of a human chromosome; or that the universe is expanding; or the reasons birds can fly; or the circulation of the blood; or the microbial cause of infectious disease. For this is science, the accurate unraveling and description of the world as it is. Instead, as further conversation revealed, he was disappointed in the ways in which science

has been used.

He was disappointed that scientific progress has not cured the world's ills by abolishing wars and starvation; that gross human inequality is still widespread; that happiness is not universal. My friend made a common mistake — a basic misunderstanding in the nature of knowledge. Knowledge is amoral — not immoral but morality neutral. It can be used for any purpose, but many people assume it will be used to further their favorite hopes for society — and at this is the fundamental flaw. Knowledge of the world is one thing; its uses create a separate issue. To be disappointed that our progress in understanding has not remedied the social ills of the world is a legitimate view, but _____.

ⓐ To argue that knowledge is not progressing because of the African or Middle Eastern conflicts misses the point. There is nothing inherent in knowledge that dictates any specific social or moral application. Some people, however, are actually disappointed in learning about our universe, as contradictory as this may sound. They have a vested interest, usually quasi-religious, in doctrinaire pronouncements on the nature of the world. After all, science can be challenging to static views as it uncovers new evidence on one topic after another, deflating strongly held opinions that are not backed by fact. Obvious examples would be the few remaining people who refuse to accept the changing configuration of our Earth and its position within our galaxy, or the existence of other galaxies. They are mired in outdated assumptions that lead them into an untenable position. A famous historic example is Galileo's astounding pronouncement that the Earth orbits around the sun rather than the sun revolving around the Earth. This contradicted the Catholic Church's proclamation on the centrality of our small, inhabited sphere. Galileo was commanded to recant observable

evidence, and he was placed under house arrest. But, as his commented, this recantation did not affect the motion of the Earth. The discovery of the evolution of new species is morally reprehensible to those whose religion teaches a contrary doctrine. Again, people who accept unupportable decrees are emotionally upset by fact. The exact age of the universe is not known, but evidence moves it into the remote past — perhaps to the twelve-thirteen billion years ago era. This longevity provokes outrage from disciples of religious systems who are taught some recent arbitrary figure.

3-1. 빈칸에 들어갈 말로 가장 적절한 것을 고르세요.

- ① to confuse this with the progress of knowledge is absurd
- ② to know the nature of knowledge is to practice its moral value
- ③ to remove social inequality is the inherent purpose of knowledge
- ④ to accumulate knowledge is to enhance its social application
- ⑤ to make science progress is to make it cure social ills

3-2. 밑줄 친 @this 가 가리키는 것은?

.....

3-3. 이 글의 작가가 ㉔ “To argue that knowledge is not progressing ... misses the point” 와 같이 주장하는 이유를 설명하세요.

.....

3-4. 위 글을 50자 내외의 우리말로 요약하세요.

.....

■ 자기평가 설문

(1번 문항에 대한 설문에 답해주세요.)

Survey Questions	전혀 아니다 ←→ 매우 그렇다				
1. 나는 지문의 내용에 친숙하다. (지문을 미리 읽어 보거나 공부한 적이 있음)	1	2	3	4	5
2. 나는 내가 선택한 답이 정답이라고 확신한다.	1	2	3	4	5
3. 위의 읽기 지문은 이해하기 쉽다.	1	2	3	4	5
4. 지문에 주어진 정보가 ‘빈 칸 채우기 문제’를 푸는데 충분하다고 생각한다.	1	2	3	4	5

4.

Long before Walt Whitman wrote *Leaves of Grass*, poets had addressed themselves to fame. Horace, Petrarch, Shakespeare, Milton, and Keats all hoped that poetic greatness would grant them a kind of earthly immortality. Whitman held a similar faith that for centuries the world would value his poems. But to this ancient desire to live forever on the page, he added ①a new sense of fame. Readers would not simply attend to the poet's work; they would be attracted to the greatness of his personality. They would see in his poems a vibrant cultural performance, an individual springing from the book with tremendous charisma and appeal. Out of the political rallies and electoral parades that

marked Jacksonian America, Whitman defined poetic fame in relation to the crowd. Other poets might look for their inspiration from the goddess of poetry. Whitman's poet sought_____. In the instability of American democracy, fame would be dependent on celebrity, on the degree to which ⑥the people rejoiced in the poet and his work.

4-1. 빈칸에 들어갈 말로 가장 적절한 것을 고르세요.

- ① poetic purity out of political chaos
- ② fame with political celebrities
- ③ a refuge from public attention
- ④ immortality in literature itself
- ⑤ the approval of his contemporaries

4-2. 밑줄 친 ⑥a new sense of fame 을 설명하세요.

.....

4-3. Whitman과 그의 시에 대해서 왜 대중이 ⑥와 같이 반응했는지 설명하세요.

.....

4-4. 위 글을 50자 내외의 우리말로 요약하세요.

.....

▣ 자기평가 설문

(1번 문항에 대한 설문에 답해주세요.)

Survey Questions	전혀 아니다 ←				→ 매우 그렇다
1. 나는 지문의 내용에 친숙하다. (지문을 미리 읽어 보거나 공부한 적이 있음)	1	2	3	4	5
2. 나는 내가 선택한 답이 정답이라고 확신한다.	1	2	3	4	5
3. 위의 읽기 지문은 이해하기 쉽다.	1	2	3	4	5
4. 지문에 주어진 정보가 '빈 칸 채우기 문제'를 푸는데 충분하다고 생각한다.	1	2	3	4	5

국 문 초 록

언어 평가에 있어서 수험자들의 언어 능력에 영향을 줄 수 있는 요인에 대한 논의가 있어 왔다. 그 중에서도, 한국인 EFL 학습자들의 영어 읽기 수행에 가장 많은 영향을 미치고 많이 연구되는 분야는 수험자가 받는 입력의 특성, 특히 영어 읽기 시험에서 지문의 길이이다. 그러나 한국인 영어 학습자가 받는 입력의 특성 중 지문의 길이를 주 변인으로 한 선행연구들은 그들의 영어 읽기 과정에 대한 설득력 있는 논거를 제시하지 못한다. 따라서, 본 연구는 표준화시험에서 쓰이는 긴 원문에서 발췌한 짧은 지문이 영어를 외국어로 학습하는 한국인 EFL 학생들의 영어 텍스트 이해력에 주는 영향을 고찰한다. 또한, 서로 다른 길이(짧은 대목을 발췌한 글과 원래 길이의 글)의 영어 독해 지문을 이해하기 위해서 학습자들이 사용하는 읽기 전략과 추론적 사고 형성 과정을 비교하고 상세히 기술하고자 한다.

연구참여자는 한국에서 영어를 외국어로 학습하는 한국인 대학생 100명, 고등학생 102명으로 총 202명이 참여하였다. 실험에 쓰이는 문제지는 A유형과 B유형 두 가지이며, 각 유형은 짧은 지문 2개, 긴 지문 2개 총, 독해 지문 4개와 각 지문에 상응하는 영어 독해 문제들로 구성되어 있다. 짧은 지문은 한국 상황의 표준화시험에서 쓰이는 지문의 길이와 유사한 평균 150 단어의 지문으로 구성되어 있으며, 긴 지문은 짧은 지문이 발췌된 원문으로 구성되어 있고 길이는 짧은 지문보다 2~3배 정도 길다. 첫번째 단계에 참여한 100명중 각 능숙도 별로 2명, 총 4명의 학생들을 선정하여 각 숙달도 집단 별로 짧은 지문과 긴 지문을 읽고 각 영어 문장을 하나씩 읽고 해석하면서 떠오르는 모든 생각을 구술하도록 하였다.

짧은 지문과 긴 지문을 읽었을 때 상, 중, 하 집단의 대학생들의 영어 수행 결과는 전체적으로 짧은 지문보다 긴 지문에서 학생들의 영어 점수가 통계적으로 유의미하게 높은 것으로 나타났다. 또한, 문제 유형별로 지문의 길이에 따른 학생들의 영어 읽기 능력의 차이가 나타나는 것을 확인하였다. 즉, 학생들의 추론적 사고를 요하는 주관식 문항의 경우 짧은 지문보다 긴 지문에서 학생들의 영어 점수가 통계적으로 유의미하게 높은 것으로 나타난 반면, 객관식 빈칸 채우기 문항의 경우, 지문 길이에 따른 학생들의 읽기 점수 차이는 통계적으로 유의미하지는 않았다. 결론적으로, 학생들의 실제 영어 읽기 과정과 더욱 유사한 주관식 문항의 평가 방식에서 지문이 더 많이 주어지는 것이 학습자들의 영어 읽기 수행을 돕는 것으로 해석할 수 있다. 고등학생들도 대학생들의 결과와 유사하였는데, 능숙도 집단 별로 객관식, 주관식 문항의 점수의 결과는 상, 중, 하 집단 모두 객관식, 주관식 문항에서 짧은 지문보다 긴 지문을 읽고 풀었을 때 점수가 다소 높은 것으로 보였다. 주관식 문항의 경우, 상 집단에서 지문 길이가 학생들의 영어 읽기 수행에 통계적으로 유의미한 영향을 주는 것으로 나타난 반면, 객관식 문항(빈 칸 채우기)의 경우에는 짧은 지문과 긴 지문에서의 영어 읽기 수행 능력 차이가 거의 없었다. 즉, 지문이 더 많이 주어지는 것이 빈 칸 채우기 문제를 풀기를 위한 학생들의 수행에는 별다른 영향을 주지 못했지만, 지문을 제대로 이해하기 위한 추론 능력을 측정하는 개방형 문제를 푸는 데는 통계적으로 유의미한 영향을 끼친 것으로 보인다.

지문 길이에 따른 대학생과 고등학생의 인식을 파악하기 위해서 세 가지 문항으로 구성된 자기평가 설문을 각 지문을 읽은 직후에 답하게 하였고, 지문을 읽고 느끼는 인식 면에서는 대학생과 고등학생이 상반된 결과를 보였다.

대학생들은 긴 지문이 짧은 지문보다 더 쉽다고 느꼈고, 짧은 지문보다 긴 지문을 읽고 자신의 답에 더 확신을 가졌으며, 문제를 풀기 위해서 주어진 정보가 긴 지문이 훨씬 더 충분하다고 느꼈다. 이와 반대로, 고등학생들의 경우는 실제 영어 읽기 수행 결과에서 짧은 지문보다 긴 지문에서 다소 점수가 높았음에도 불구하고 본인들 스스로는 짧은 지문보다 긴 지문에서 본인들의 영어 실력을 확신하지 못하고, 지문이 더 어렵다고 느꼈으며, 지문이 더 길었음에도 불구하고 문제를 풀기 위한 충분한 정보가 주어지지 못했다고 인식했다. 본 연구의 결과와 교육적 함의가 결론에서 심층적으로 논의된다.

주요어: 영어 독해 능력, 표준화영어시험, 지문 길이, 문제 유형, 추론

학 번: 2015-21849