2018

The Role Of Cohesion In Second Language Reading Comprehension

Alisha Biler
*University of South Carolina*

Follow this and additional works at: [https://scholarcommons.sc.edu/etd](https://scholarcommons.sc.edu/etd)

Part of the [Linguistics Commons](https://scholarcommons.sc.edu/etd)

**Recommended Citation**


This Open Access Dissertation is brought to you by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact dillarda@mailbox.sc.edu.
THE ROLE OF COHESION IN SECOND LANGUAGE READING COMPREHENSION

by

Alisha Biler

Bachelor of Science
Bowling Green State University, 2007

Master of Arts
University of South Carolina, 2010

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Linguistics

College of Arts and Sciences

University of South Carolina

2018

Accepted by:

Nina Moreno, Major Professor
Robin Morris, Committee Member
Mila Tasseva-Kurktchieva, Committee Member
Danielle McNamara, Committee Member
Cheryl L. Addy, Vice Provost and Dean of the Graduate School
ACKNOWLEDGEMENTS

There are a number of people to whom I owe the deepest of gratitude in the completion of this dissertation. Without the support of professors, colleagues, family, and friends, this project would never have reached fruition.

First, I would like to thank my director, Dr. Nina Moreno, for her steadfast support and encouragement throughout the entirety of the process. From the very beginning, her expertise has guided this project. When I proposed absurd deadlines along with a seemingly unending array of questions and dilemmas, she was unfailing in offering advice and guidance. My project would not be what it is today without her support. I also owe a deep gratitude to Dr. Robin Morris, who challenged me to ask deeper questions and fully develop this project into something I am proud to present today. I am grateful for the time she invested in sharing her knowledge and extensive experience in reading research, without which I could not present this paper.

Dr. Mila Tasseva-Kurktchiev deserves a medal for the number of times she saves the day for needy Linguistics students, myself more so than most. I am forever indebted to her encouragement and tactical support from the beginning of my time in the MA program until its completion. Thank you also to Dr. Danielle McNamara whose research inspired this study and whose feedback and advice helped me to refine the design of this study.
This research would not have been possible without the financial support of The International Research Foundation for Language Learning/Cambridge English Assessment as well as the Linguistics program at the University of South Carolina. Many thanks also to Keiko Bridwell for assistance in data cleaning, Jenn Olejarczyk for teaching me everything I could ever hope to know about running an eye-tracking study, as well as Wilma Sims and Chris Farina for statistical counseling.

To my colleagues at the English programs for Internationals, thank you for encouraging your students to participate in my study and for encouraging me daily throughout this journey. Most outsiders advised me that working a full-time job while writing a dissertation was a regrettable choice, but I wouldn't change a single day that I got to work with any of you. I also extend my gratitude towards my peers in the Linguistics program for helping me to make it to the other side of this graduate school journey.

There are a score of others who impacted the work within this document, and for each person, I am endlessly grateful.
ABSTRACT

Reading in a second language (L2) is a critical aspect of language acquisition, yet gaps remain in the literature regarding the extent to which textual factors impact reading difficulty. There is consensus that complex vocabulary and grammar affect L2 comprehension (Koda, 2005), and this is evidenced through the numerous traditional readability formulas, such as Flesch-Kincaid (FKGL). However, critics argue that discourse-level features, such as cohesion, also impact reading difficulty and must be included in difficulty analyses (Carrell, 1987).

One aspect of cohesion is content word overlap, or the number of content words repeated in a text; this measure is included in Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004), a recent readability tool. Content word overlap (CWO) is thought to facilitate reading by making explicit text connections to relieve the processing burden (Kintsch & Van Dijk, 1978). This claim is supported in the L1 literature (Britton & Gulgoz, 1991), but is relatively unexplored in L2 contexts.

This study utilized two experiments to capture effects of CWO in L2 reading comprehension. Experiment one looked at L2 reading comprehension scores (N=1131 scores on 18 tests), and compared the results to predictions of text difficulty made by two divergent readability tools. Results indicated that Coh-Metrix, which measures CWO, was more reliable in predicting comprehension than FKGL, which considered lexical/grammatical complexity alone.
Experiment two utilized eye-tracking measurements of L1 (n=31) and L2 readers (n=63) in passages with varying amounts of CWO of a target word. Results indicated that all nonnative speakers were susceptible to longer target word reading times when no overlap was present ($p < .05$). Additionally, lower proficiency L2 learners benefitted from more overlap than higher proficiency learners ($p < .05$) in target reading time and integrating the target into the surrounding text.

The results of both experiments support the role that content word overlap affects second language reading in both overall comprehension and localized processing. Thus, when considering models of L2 reading, it is necessary to include cohesion in addition to lexical and grammatical complexity. These findings have implications for pedagogues, and assessment developers to ensure that reading difficulty is accurately measured.
TABLE OF CONTENTS

Acknowledgements........................................................................................................ iii
Abstract.............................................................................................................................. v
List of Tables ...................................................................................................................... ix
List of Figures ..................................................................................................................... xi
List of Abbreviations ....................................................................................................... xii

Chapter 1: Introduction .................................................................................................... 1
  1.1 Statement of the Problem ......................................................................................... 1
  1.2 Definition of Terms ................................................................................................. 3
  1.3 Dissertation Overview ............................................................................................ 4

Chapter 2: Literature Review ........................................................................................... 5
  2.1 Cognitive Processes in Reading ................................................................................ 5
  2.2 Assessing Reading Comprehension ......................................................................... 17
  2.3 Rationale for Current Study and Research Questions ............................................. 25

Chapter 3: Methodology ................................................................................................. 29
  3.1 Experiment One ....................................................................................................... 29
  3.2 Experiment Two ....................................................................................................... 35

Chapter 4: Results .......................................................................................................... 50
  4.1 Experiment One Results ......................................................................................... 50
  4.2 Experiment Two Results ......................................................................................... 60

Chapter 5: Discussion ..................................................................................................... 76
5.1 Experiment One Discussion ................................................................. 76
5.2 Experiment Two Discussion ............................................................... 82
5.3 General Discussion ............................................................................. 91
5.4 Limitations and Future Directions ....................................................... 93
5.5 Pedagogical Implications .................................................................. 97
5.6 Conclusions ...................................................................................... 99
References .............................................................................................. 100
Appendix A: Experiment one sample reading assessment test (6b) ............. 109
Appendix B: Experiment two passage metrics and familiarity ratings ........ 113
Appendix C: Experiment two passages in overlap conditions .................. 116
LIST OF TABLES

Table 3.1 Test titles, word counts, and administration dates for assessments in experiment one .................................................................32
Table 3.2 Total reading comprehension scores collected ...........................................33
Table 3.3 Summary of experiment one participant language background .................33
Table 3.4 Summary of participants in experiment two..............................................38
Table 3.5 Summary of passage familiarity ratings for L1 and L2 participants (Range 1-6) ......................................................................................40
Table 3.6 Content word overlap in adjacent sentences by passage proficiency level ......42
Table 3.7 Syntactic difficulty measurements by passage proficiency level ...............43
Table 3.8 CELEX frequency of content word norms (McNamara et al., 2014) and means by passage proficiency level .........................................................44
Table 3.9 RDL2 norms (McNamara et al., 2014) and means by passage proficiency level .................................................................................................44
Table 3.10 Summary of passage reading by experimental condition and difficulty ......47
Table 4.1 Number of participants and mean proficiency scores by level and test.........51
Table 4.2 Mean comprehension scores (%) by tests and levels..................................52
Table 4.3 Results of ANOVA for comprehension differences on tests within a level ......53
Table 4.4 Readability scores for each test by Coh-Metrix (RDL2) and Flesch-Kincaid(FKGL) ........................................................................................................55
Table 4.5 Test-pair differences in RDL2 readability predictions alongside significance values for actual comprehension score differences. .........................................................56
Table 4.6 Test-pair differences in RDL2 readability predictions alongside significance values for actual comprehension score differences. .........................................................59
Table 4.7 Overall reading measures by group and passage difficulty .........................63

Table 4.8 Mean gaze durations (msec) on target words by passage difficulty ...............67
LIST OF FIGURES

Figure 3.1 Example passage illustrating locations of critical and target words...............41
Figure 3.2 One example passage presented in all four overlap conditions.....................46
Figure 4.1 Example of gaze duration............................................................................66
Figure 4.2 Example of regression path duration.............................................................66
Figure 4.3. L1 participant example passage with eye-movements, no overlap ............68
Figure 4.4. High L2 participant example passage with eye-movements, all overlap ......69
Figure 4.5 Low L2 participant example passage with eye-movements, distant overlap ...69
Figure 4.6 High L2 participant example passage with eye-movements, close overlap .....70
Figure 4.7: Mean gaze duration times by overlap and proficiency group .....................71
Figure 4.8 Low L2 target reading time (msec) by passage overlap and difficulty ............73
Figure 4.9 Regression path duration from the target word by group and difficulty .........74
Figure 4.10 Regression path duration from the target word by group and overlap ........74
Figure 5.1 Graph of RDL2 predictions by level; tests with significant comprehension differences noted with *.................................................................78
Figure 5.2 Graph of FKGL predictions by level; tests with significant comprehension differences noted with *.................................................................79
LIST OF ABBREVIATIONS

CEFR ................................................. Common European Framework of Reference
EFL .......................................................... English as a Foreign Language
EPI.......................................................... English Programs for Internationals
FKGL ....................................................... Flesch-Kincaid Grade Level Index
L1 .................................................................. First language
L2 .................................................................. Second Language
RDL2 .......................................................... Coh-Metrix L2 Readability
SLA ............................................................. Second Language Acquisition
VST ............................................................. Vocabulary Size Test
CHAPTER 1
INTRODUCTION

1.1. Statement of the Problem

The task of assessing the difficulty, or readability, of a text for first language (L1) readers has been a pursuit of researchers for almost a century, beginning with Lively and Pressey (1923). For educators, the ability to clearly categorize the level of a text is essential for matching learners to appropriate texts to maximize learning. Textbook writers, materials developers, and assessment writers need to be aware of how to construct texts carefully so that meaning can be understood clearly. Reading is foundational to the learning process, both language learning and general learning, and as such, the need for readable texts cannot be understated.

While the significance of identifying text difficulty is generally acknowledged, methods for objectively determining and quantifying the difficulty of a given text are contested. Traditionally, the U.S. education system has relied upon readability formulas measuring word and sentence difficulty as the primary methods for assessing L1 text difficulty (e.g., Flesch-Kincaid Grade Level Index: Kincaid, Fishburne, Rogers, & Chissom, 1975; The Lexile Framework: Stenner, 1996). However, critics have argued that such measures only evaluate the surface reading characteristics and ignore the deeper psychological processes involved in reading (Carrell, 1987; Davison & Kantor, 1982; Graesser, McNamara, & Kulikowich, 2011).
Other researchers have proposed the structure of a text greatly impacts comprehension difficulty. A text with fewer conceptual gaps in its organization, or a highly coherent text, has been found to aid L1 readers’ comprehension (Beck, McKeown, Sinatray, & Loxterman, 1991; Britton & Gülgoz, 1991). A number of studies have given evidence for the inclusion of cohesion in determining the difficulty of a text for L1 readers, especially readers with minimal knowledge of the topic of the text (McNamara, Kintsch, Butler-Songer, & Kintsch, 1996).

For the second language (L2) learner, the challenge of ascertaining a text's difficulty is equally important. In terms of learning a second language, errors made in productive skills, like grammatical errors in writing or pronunciation errors in speaking, are easily diagnosed and corrected. However, when learners encounter comprehension difficulties in reading, there is no simple method for isolating the source of the breakdown; it is likely a combination of multiple factors at both the sentential, textual, and contextual level. Thus, readability metrics used in L1 reading, which assess only sentence level difficulty, are ineffective at predicting L2 reader difficulty despite their frequent usage by L2 educators and materials developers.

The purpose of this research is to investigate potential causes of L2 reading comprehension difficulty, namely lexical difficulty, syntactic difficulty, and content word overlap (cohesion) through two complementary experiments. The first experiment of the study looks broadly at readability analysis tools and their efficacy at predicting L2 reader comprehension. By directly comparing the predictions of Flesch-Kincaid Grade Level Index (which measures only lexical and syntactic difficulty) with Coh-Metrix L2 Readability Index (which includes content word overlap measurements along with lexical
and syntactic difficulty) on L2 reading comprehension assessments, the factors underlying reading comprehension difficulty will become more clear. The second experiment takes a more in-depth look at the extent of the effects cohesion has on comprehension. This experiment controls both lexical difficulty and content word overlap in an online reading task to determine in real-time processing how these measures affect reading. While post-reading comprehension questions can provide a glimpse into the mind of the reader, the measurements of eye-movement patterns will be more beneficial in understanding what makes reading difficult as it measures processing in real time. Thus, the off-line nature of experiment one is complemented by the on-line task in experiment two to provide a more comprehensive analysis of L2 reading text difficulty.

1.2 Definition of Terms

Content word overlap: The repetition of content words throughout a given text (Crossley, Greenfield, & McNamara, 2008). Used in the L2 readability formula of Coh-Metrix as the repetition of content words occurring in adjacent sentences. Used in this dissertation as the measure of cohesion.

Gaze duration: A type of eye-movement measure that is the sum of all fixations in a region during the first reading of the region. This term is also known as first-pass reading time and is indicative of "early" reading processing, e.g. word identification (Clifton, Staub, and Rayner, 2007).

Readability: A term used to designate the measured ease of understanding a text. In the context of readability formulas, this is used almost exclusively for text-based
characteristics, such as word or sentence length, and ignore sources of reading difficulty attributed to the reader, such as motivation or background knowledge.

Regression path duration: A type of eye movement measure that includes the summed fixation time of the first fixations on a region of interest plus any regressive fixations made prior to moving to the right of the region of interest. Also referred to as go-past time. This measure is typically considered a "late" processing measure, indicative of text integration (Rayner, Sereno, Morris, Schmauder, & Clifton, 1989).

1.3 Dissertation Overview

This dissertation is organized into five chapters. The present chapter introduces the topic and states the operationalized definition of the potentially unfamiliar or ambiguous terms that will be used throughout the manuscript. Chapter two reviews the relevant literature on the nature of reading including the unique challenges of L2 reading, the role cohesion in L1 reading, and how reading is assessed, namely through readability formulas and on-line measurements. The gaps in the literature are summarized and the rationale for two experiments is provided along with the research questions. The third chapter presents the research design for the two complementary experiments that seek to resolve the gaps introduced in chapter two. Chapter four presents the results of both experiments. The final chapter, chapter five, discusses the relevance of the results of each experiment and what the collective results imply for cohesion in L2 reading. It also presents limitations of the current study, suggestions for future research, and relevant applications for practitioners.
CHAPTER 2
LITERATURE REVIEW

2.1 Cognitive Processes in Reading

**General Reading Comprehension**

The question at the core of this dissertation is whether or not cohesion affects the readability of a text for second language learners. Before discussing what makes texts readable, however, it is important to obtain a broad picture of reading comprehension from a cognitive point of view and define what processes are engaged while reading. This section does not seek to present all established facets of reading that have been demonstrated empirically, as such a list is a dissertation in itself, but rather to provide a clear, basic foundation of what is entailed in reading.

Reading researchers are in agreement that at its simplest definition, reading is a multi-step, effortful process involving rapid linguistic processing, at the word, sentence, and text levels (Rayner, Pollatsek, Ashby, & Clifton, 2012). As such, reading comprehension is not a simple construct that can be easily quantified. Johnston (1983) defines comprehension as follows:

Reading comprehension is considered to be a complex behavior which involves conscious and unconscious use of various strategies, including problem-solving...
strategies, to build a model of the meaning which the writer is assumed to have intended. Most of this model must be inferred, since text can never be fully explicit and, in general, very little of it is explicit because even the appropriate intensional and extensional meanings of words must be inferred from their context. (p. 17)

Thus, there is no singular theory of reading as reading is not a singular skill but rather a series of intricate processes that coordinate simultaneously, utilizing cognitive resources.

Decades of reading research have demonstrated the necessity of a variety of skills in fluent reading. Although the studies mentioned here are based on L1 reading in English, they are also the foundation of L2 reading (to be discussed more fully in the next section). The skills necessary for fluent successful reading can be divided into lower- and higher-level processes.

Lower-level processes include skills at the word and sentence level. Upon encountering a word in a text, a reader must begin with rapid word recognition. Successful word recognition includes strong grapheme-sound correspondence (Rayner et al., 2012), and rapid lexical access to both meaning and syntactic properties of the word (for review, see Grabe, 2009). Simultaneous sentence parsing also occurs while reading. Readers fill in their understanding of the sentence by accessing word meaning and assigning that meaning to a role in the sentence during reading (Perfetti & Adlof, 2012).

In addition to processes at the word and sentence level, higher-order processes at the discourse level are required for fluent reading. Higher-order processes are defined as those 'closer to conscious introspection on the part of the reader' (Grabe, 2014, p.10). The
information obtained from clause-level processing is integrated to build a discourse-model of the text based on linking information throughout the text. This basic comprehension of the text is combined with the reader's background knowledge, inferencing, and the reader's attitudes to have an overall interpretation of the text (Kintsch, 2012).

A multitude of other factors are also significant in reading. These include, but are not limited to, text genre, reader motivation, as well as purpose and context of reading. Additionally, individual differences such as working memory and attention capacity are also at play. In sum, a variety of reading skills are working in concert to achieve reading comprehension of a given text. For the purposes of this dissertation, the role of word and sentence level processing as well as text integration are of primary investigation as they pertain to second language reading.

L2 Reading Comprehension

When comparing adult L2 reading to L1 reading, the processes involved in comprehension are increasingly complex. L2 reading entails all of the processes described in the above section, but adult L2 readers have the additional complication of two competing linguistic systems to manage. In word recognition, bilingual readers must suppress word recognition in the L1 in order to activate the L2 lexicon, which has been shown to increase reading time (Kern, 1994; Koda, 2005). The same is also true in managing the orthographic, syntactic, and stylistic textual cues of the L1 system. However, having an existing L1 system is not exclusively a hindrance to L2 reading comprehension. Beginning adult L2 readers have existing advanced cognitive processing
and reasoning skills as well as reading strategies that they have carried over from their L1 (Dressler & Kamil, 2006; Koda, 2005). As such, even developing L2 readers have additional cognitive abilities beyond their L2 proficiency. However, researchers disagree as to when and how L1 reading strategies transfer to L2 reading, and this debate will be addressed below.

Alderson's (1984) seminal question has continued to echo throughout the literature regarding L2 reading processing: "Reading in a foreign language: A reading problem or a language problem?" The literature as summarized in the subsequent paragraphs points to a model where L2 knowledge plays a greater role in reading processing for beginning level learners than it does for advanced learners. After a learner has reached a certain threshold in L2 proficiency, L1 reading abilities can begin to have an effect, known as the threshold hypothesis.

The threshold hypothesis (Clarke, 1980) holds that readers are capable of transferring L1 reading skills into their L2 reading but only after they have reached an undefined threshold of L2 proficiency. Clarke states that a reader's limited control in the L2 interferes or inhibits the good reading strategies of the learner. He writes:

Perhaps there are not 'good readers' and 'poor readers' but merely 'good' and 'poor' reading behaviors which characterize most readers at different times; when one is confronted with difficult reading (whether because of complex language or unfamiliar content) one is likely to revert to poor reading behaviors. (p. 206)
In other words, even 'good L1 readers' will appear like 'poor readers' in an L2 context until they can reach an undefined level of L2 proficiency. They will exhibit 'poor reading behaviors' as they are unable to access their L1 'good reading behaviors.'

The threshold hypothesis has been supported in a number of studies summarized below which uniquely compare L1 and L2 reading within the same individual. In previous studies, comparisons between L1 readers and L2 readers were ineffective in addressing this issue because there were a number of individual differences (working memory capacity, cognitive processing skills, etc...) that could interfere (Alderson, 1984). However, in studies which compare the L1 and L2 reading abilities of the same individual, these differences are held constant and more can be understood about L2 reading.

Carrell (1991) compared L1/L2 learners of Spanish/English on their abilities to answer inferencing comprehension questions on both L1 and L2 texts. She found that L1 English participants' scores on L1 comprehension tests only had a significant correlation with scores in the L2 comprehension tests for advanced learners, meaning being a good reader in the L1 implied being a good reader in the L2. For beginning level L2 learners, a high L1 comprehension score did not affect L2 scores, implying that low-level L2 learners did not transfer their 'good' L1 reading strategies.

Lee and Schallert (1997) looked at 800 Korean learners of English and found evidence that while L2 proficiency played a significant role for all learners in L2 reading comprehension, L1 reading ability had a greater influence for the more advanced English learners. L2 proficiency level was measured by a vocabulary assessment and a
grammatical judgment task; all comprehension questions in the L1 and L2 reading assessment were inference-based questions. Similar to Carrell (1991), the authors concluded that L2 beginners cannot transfer their L1 reading skills until they achieve a certain level of L2 proficiency.

Similarly, Bossers (1991) compared L1 Turkish learners of Dutch of varying proficiencies. When comparing all participants as a group, L2 knowledge, as measured by vocabulary and grammatical knowledge, was the strongest contributing factor in reading comprehension; however, when participants were separated by L2 proficiency, L1 reading ability became the most significant contributing factor for the advanced learner group only. Thus, these studies seem to support the claim of the threshold hypothesis that L1 reading abilities transfer after L2 learners have crossed a certain threshold of L2 proficiency.

Other researchers have criticized the premises of the threshold hypothesis. One critique is that the threshold hypothesis fails to define precisely what the 'threshold' is, and how we know when students have crossed such a threshold (Hulstijn, 1991). Other critiques claim that even the most advanced L2 learners still exhibit differences when reading in their L2. Advanced L2 readers may perform similarly on both L1 and L2 comprehension tests yet under certain conditions, they are found to read about 30% slower in the L2 (Coady, 1979; Segalowitz, Poulsen, Komeda, 1991). Therefore, the source of differences between L1 and L2 readers cannot be L2 proficiency knowledge alone.
Additionally, the types of instruments used in the threshold experiments need to be examined. For all of the studies listed above in support of the threshold hypothesis, the measure of L2 proficiency was a measure of lexical and grammatical knowledge. Yet, many of these studies do not report the measures by which they controlled reading difficulty of the L2 texts. When they are reported, the measures used were traditional readability formulas, such as the Dale-Chall readability formula, which rely exclusively on lexical and syntactic measures of the text. As such, it is not surprising that vocabulary and grammatical proficiency levels correlated with texts defined as having complex vocabulary and grammar. Although lexical and grammatical knowledge is an important predictor in L2 reading, the evidence presented in the above studies is confounded by the means of measuring text difficulty, and other contextual measures need investigation.

Research has shown that less skilled readers may rely more heavily on contextual information to compensate for the lack of automaticity in word recognition (Coady, 1979). If texts used in the threshold hypothesis studies had more significant contextual support in the form of connectives or overlap, it is possible that low-level L2 learners might have been able to use their L1 inference strategies and improve comprehension. In other words, more thoroughly analyzed texts could indicate that low-level L2 processing, though hindered by a lack of vocabulary and grammatical knowledge, can be scaffolded with a more cohesively structured text.

What can be concluded from the threshold hypothesis studies is that when low-level learners are hindered in lexical and syntactic processing, they do not have enough processing capacity to make inferences while reading as they might be able to in their L1. This may be compensated by increased cohesive devices in the text, but the literature
does not report on these textual aspects, leaving this question unanswered. The results of these studies and the participants' inability to answer inference questions would be better informed by analyzing the organization and cohesion of the texts rather than only focusing on lexical and syntactic complexity. More authentic means of assessing comprehension are needed beyond the surface features of lexical and syntactic complexity of texts, including discourse-level and text integration aspects such as cohesion.

**Cohesion**

Due to the significant role of discourse integration in reading comprehension, a number of studies have examined the role of cohesion in reading comprehension. As yet, there are minimal studies investigating cohesion in L2 reading, so the majority of this discussion will focus on cohesion in L1 reading.

Cohesion itself is a broad term and in need of further specification. Various manipulations of cohesion found in the literature include referential cohesion (Gernsbacher & Robertson, 2002), temporal cohesion (Ohtsuka & Brewer, 1992), and causal cohesion (Myers, Shinjo, & Duffy, 1987). While each type of cohesion plays a role in processing, one of the most significant influences based on Kintsch and van Dijk's (1978) model of text comprehension is co-referentiality and its benefits for working memory limitations.

According to Kintsch and van Dijk's (1978) model, comprehension of discourse can be represented at three levels: 1) *surface representation*, the understanding of the words and sentences as written; 2) *text base*, where the surface representations are
understood in connection with the larger discourse as a network of propositions; and 3) situation model, the knowledge learned from the text becomes integrated with the reader’s background or general knowledge. At the text base level, the authors write, “Referential coherence is probably the most important single criterion for the coherence of text bases... If a text base is found to be referentially coherent, that is, if there is some argument overlap among all of its propositions, it is accepted for further processing” (p.367). If there are no gaps between the propositions, processing is more rapid as inferencing does not need to occur in order to integrate the information either locally (between sentences) or globally (throughout the entire text). Thus, this dissertation continues to investigate cohesion as the amount of overlap between arguments.

The role of argument overlap in L1 reading has been explored in a variety of contexts. Rashotte and Torgesen (1985) found that non-fluent learning disabled L1 adolescent readers improved their reading speed from reading stories with shared vocabulary rather than multiple readings of the same story. Participants orally read unrelated texts, each containing 10-12 sentences at a grade two reading level, verified using Fry's (1977) readability formula. The authors manipulated the passages into two conditions: minimal word overlap, where no more than 20 words were common to more than three of the stories; and high overlap, where at least 60 of the words used were common to three or more of the stories. After reading the stories in one of the conditions, participants answered a comprehension question. Results indicated that students' reading speeds improved when they read texts in the word overlap condition; however, there were no gains in comprehension in the overlap condition. The authors attribute the lack of
comprehension gains to the already high comprehension scores of the readers despite their poor reading fluency.

Britton & Gül göz (1991) explored multiple approaches to cohesion in order to test Kintsch's computer program (Miller & Kintsch, 1980) which automatically looked for cohesion breaks based on Kintch & van Dijk's (1978) model of reading. Gaps in cohesion were identified as places where the reader would need to make an inference to bridge new information with the existing discourse. Using passages from *U.S. Air Power: Key to Deterrence* (U.S. Air Force Reserve Officers' Training Corps, 1985), the authors wrote revised versions based on three cohesion principles: 1) argument overlap in adjacent sentences; 2) placing old information before new information in sentences; and 3) making any implicit references explicit. An example of argument overlap modification is listed below:

1a) Original title and first sentence (*emphasis added*):

   Air War in the North, 1965

   By the fall of 1964, Americans in both Saigon and Washington had begun to focus on Hanoi.

1b) Rewritten title and first sentence (*emphasis added*)

   Air War in North Vietnam, 1965

   By the beginning of 1965, Americans in both South Vietnam and the U.S. had begun to focus on North Vietnam. (p.345)
L1 university participants read one of the texts and either took a free-recall or multiple-choice comprehension test. Participants reading the more cohesive version, comparable to 1b described above, scored significantly higher on the free-recall comprehension measure, but not on the multiple-choice test. Additionally, participants reading the more cohesive version read the text faster than the non-modified version; however, the results were not significant.

McNamara et al., (1996) partially replicated Britton & Gülgööz (1991) by exploring the effects of local and global cohesion on middle school L1 readers with academic texts. The original text was modified by the authors for four types of cohesion: 1) replacing ambiguous pronouns with explicit noun phrases (e.g., replacing *it* with *the heart*); 2) including elaborative links between unknown ideas and known ideas (e.g., "This disease usually follows a sore throat caused by bacteria known as streptococci. *This is often called strep throat.*"); 3) increasing connectives between sentences (e.g., *however, because*); and 4) replacing words to add more argument overlap between sentences (e.g., replacing *person* and *cases* with *baby/babies*). (p.21)

Participants' prior knowledge was assessed with a sorting task and a prior-knowledge questionnaire. Participants then completed a second sorting task as well as answered text-based questions after reading one version of the text. For participants with low-background knowledge, the maximally cohesive text resulted in greater comprehension scores on questions of all types, as well as better performance in the sorting task. For high-knowledge readers, however, the minimally cohesive text resulted in better sorting scores as well as inference-based questions. The authors posit that the minimally
cohesive text requires the readers to make more inferences, resulting in greater active processing. Deeper processing occurs for the high-knowledge readers when they must make the inferences on their own; however, low-knowledge readers do not have the background knowledge to make such inferences and benefit from explicit connections being made for them. Thus, cohesion does not affect all readers in the same way.

Vidal-Abarca, Martinez, & Gilabert (2000) found that passages rewritten only with argument overlap changes found an improved performance on delayed recall tasks, but no significant improvement in immediate comprehension. The study also used middle school L1 readers with modified expository texts. Following Britton & Gülgöz (1991), they repaired cohesion breaks with argument overlap, and assessed participants on immediate recall, inference testing, and delayed recall. The argument overlap condition showed no significant effects for immediate recall or inferences, but did show improvement over the original text version for delayed recall.

While the literature on cohesion in L1 reading is extensive, there is little research on the effects of cohesion for L2 readers. What can be concluded based on the previously described L2 research is that learners struggling with word and sentential-level understanding due to limited processing resources may also have difficulties with inference-making requirements of low-cohesion texts. Given that only highly proficient L2 readers can make use of their L1 reading ability, it could be argued that low-proficiency L2 readers are more likely to need argument overlap to assist with inference-making while higher proficiency learners may be able to build inferences without such textual support. Increasing cohesion may thus alleviate the processing burden of low-
proficiency readers as it requires fewer resources. However, these effects remain to be
tested in the L2 literature.

The literature summarized above is a broad overview of the reading process and
how word and sentence level features impact L2 reading. The literature is growing on
how text-level features such as content word overlap affect L1 reading, but there are few
studies measuring cohesion directly in L2 reading. L2 studies have approached cohesion
from a readability perspective and will be introduced below.

2.2 Assessing Reading Comprehension

This section explores the means by which researchers and educators assess
reading comprehension, namely through readability tools and on-line processing
measurements. The research discussed in the previous section lays out the processes
involved in L1 and L2 reading, and this section reviews studies that have tried to quantify
the comprehension of L1 and L2 reading. First, readability tools capturing text difficulty
are discussed followed by studies measuring reading comprehension in real time
analyses.

Readability Tools

One means of indirectly assessing reading comprehension is by using readability
metrics to determine text difficulty. By assessing the difficulty of the text, we are better
able to predict how readers will comprehend it. Since the 1920s, educators and
researchers have been using readability indices for assessing the level of text difficulty
for L1 readers (Carrell, 1987). Currently, over 200 readability formulas exist with
consistent success in native language reading. Despite the prevalence of readability tools
in L1 reading, there is little empirical research on readability tools and L2
comprehension. Of the small number of L2 studies that do exist, two tools have been primarily used and therefore will be used in this dissertation: Flesch-Kincaid Grade Level Index (Kincaid et al., 1975) and Coh-Metrix (Graesser et al., 2004).

**Flesch-Kincaid Grade Level Index.** Flesch-Kincaid Grade Level Index (FKGL) assesses text readability by analyzing two aspects of a text: syntactic complexity and semantic complexity. Syntactic complexity is measured by the number of words per sentence, and similarly, semantic difficulty is measured by word length. Traditional readability formulas such as this are built on the premise that long, complex sentences with longer, and thus low-frequency, words are more difficult for readers than shorter sentences with shorter more frequently used words.

FKGL is appealing for researchers and educators to use due to its simplistic measurement and ease of access; the tool is available to anyone with Microsoft® Office on their computer. Once a text is analyzed, FKGL provides a measurement indicating the corresponding grade level in the U.S. K-12 school system with a range from 1-12. The scores can also be interpreted as the number of years of education required to understand a text. For example, an FKGL score of 9.0 implies that the text requires a ninth-grade reading knowledge or the equivalence of nine years of schooling.

While traditional readability tools such as FKGL have been used in formal educational settings, they are not without criticism. The predominant critique of such formulas is that semantic and syntactic complexity are merely the shallow features of a text. They only measure the lower level processes of reading discussed earlier and ignore the higher level processes, such as text integration. Thus, those measurements alone
cannot predict the success of reading as they do not acknowledge the deeper processes that occur in reading comprehension (McNamara, Louwerse, McCarthy, & Graesser, 2010). For example, Carrell (1987) provides the following example of two sentences that would yield the same readability from traditional formulas:

(1) The uneven numbers are one, three, five, seven, nine, eleven, and thirteen.

(2) The squares of the absolute values of the transition amplitudes are summed.

(p.24)

While acknowledging the detriment of taking two sentences in isolation, which reduces the statistical power that traditional readability formulas rely on, Carrell argues that the above sentences are not equal in processing demands. Although FKGL would assign both sentences the same readability as they have the same number of syllables per word (1.4) and words per sentence (12), the use of more advanced grammar and unfamiliar words would render sentence (2) more difficult to read. Carrell's critique aligns with the literature discussed earlier regarding the higher level factors involved in reading, namely cohesive devices and individual reader differences in background knowledge and motivation.

Readability tools are used frequently in L2 environments despite the lack of empirical studies exploring their effectiveness in L2 comprehension. The few studies evaluating readability formulas on L2 reading are presented here. One of the first studies to explore this relationship is Hamsik (1984), who found that L2 reader comprehension scores were correlated with Flesch-Kincaid readability predictions; however, the study was limited in size with only 40 participants of heterogeneous language background.
Similarly, Greenfield (1999, 2004) used Flesch-Kinciad to compare the results of cloze test scores from 200 Japanese L2 learners of English from Bormuth (1971). The Bormuth (1971) passages contained 31 academic texts, which Greenfield turned into fifth-word deletion cloze tests. The resulting correlation for FKGL and cloze score was .85, suggesting readability tools are useful in L2 reading. However, not all readability studies have found such strong correlations. Brown (1998) looked at 2,300 Japanese L2 readers’ cloze scores on 50 randomly chosen English books. He found that the correlation between cloze score and FKGL predictions ranged from .48-.55 and concluded that traditional readability formulas were not conducive for predicting L2 reading comprehension.

A possible explanation for the lack of consistency in experimental studies regarding traditional readability formulas is that L2 learners have a paucity of vocabulary and significant cultural gaps in their background knowledge, causing them to rely more strongly on the text and its structure for comprehension. Thus, traditional readability formula predictions are likely to be even more problematic for L2 readers than L1 readers due to their exclusive focus on surface features of texts. Specifically, shorter sentences may leave out connections and relations between ideas necessary for building connections within the text, forcing readers to make inferences and placing a greater burden on the processor (Blau, 1982). In these instances, longer sentences with rich context and multiple sentence connectors (e.g., 'because' or 'secondly') facilitate L2 reading more effectively than the shortest and most simplistic syntactic structure. Thus, cohesion may play a significant role for L2 learners – a factor that is not represented in
traditional readability tools but is present in a more recently established readability formula.

**Coh-Metrix.** Coh-Metrix is a readability tool that analyzes texts using over 108 different measurements. The web tool is a single interface that accesses multiple modules including various corpora, part of speech categorizers, syntactic parsers, and statistical representations of world knowledge. It is freely available to the public and can be accessed here: http://cohmetrix.com/

The primary focus of Coh-Metrix is cohesion and coherence by analyzing over 50 types of cohesion, including: causal cohesion, a ratio of causal verbs in the text by causal particles; co-referential cohesion, measured by looking at argument overlap in adjacent sentences (local) and throughout the text (global); connectives; and latent semantic analysis, including sentence-paragraph, sentence-text, paragraph-paragraph, and paragraph-text analyses (Graesser et al., 2004). The tool also includes semantic analysis by calculating word frequency based on the COBUILD corpus, hypernym/polysemy, and type/token ratio. The web tool includes FKGL in its analysis as well as its own composite L2 readability index (Crossley, Dufty, McCarthy, & McNamara, 2007). The L2 readability index of Coh-Metrix (RDL2) analyzes three distinct aspects of a text: (1) Syntactic Complexity, as measured by the number of words per sentence; (2) Co-referentiality, as measured by content word overlap between adjacent sentences; and (3) Word Frequency, as measured by the Center for Lexical Information (CELEX) frequency scores.
Coh-Metrix has been compared directly with other readability formulas in a number of L1 reading studies involving text difficulty research with mixed results. McNamara et al. (2010) selected 12 published L1 reading studies from the field of discourse psychology that manipulated cohesion levels in texts. All of the studies used extended texts that were manipulated for high- and low-cohesion versions. The authors ran the passages from the studies through Coh-Metrix, and the measurements in Coh-Metrix were able to significantly differentiate between the high- and low-cohesion versions of texts while FKGL was unable to do so.

In terms of Coh-Metrix applications to L2 reading, Crossley, Greenfield, & McNamara (2008) replicated the study from Greenfield (1999) using the Coh-Metrix RDL2 as well as FKGL. They used the same 31 passages and cloze-test scores as Greenfield (1999) and found that RDL2 scores had a correlation of .925 (FKGL had .85). The authors concluded that RDL2 was as successful as traditional readability formulas. In the limitations section, the authors acknowledge that cloze tests are not the optimal proficiency tool as they evaluate word and sentence level understanding, similar to what FKGL measures; therefore, cloze test scores may actually be more highly correlated to FKGL (p.490).

Coh-Metrix was also found to more successfully classify texts that had been intuitively simplified for L2 learners. Crossley, Allen, and McNamara (2011) analyzed a corpus of texts that contained newspaper stories simplified using intuition alone into three different difficulty levels. These texts were then compared using three readability formulas: FKGL, Flesch Reading Ease, and RDL2. Results of statistical analyses revealed that RDL2 accurately identified the level of texts more frequently (60%
Crossley et al. (2011) concluded that intuitive text simplification is represented by the unique measures included in the RDL2 formula, making it more reliable for L2 text difficulty assessment.

In sum, the literature on the reliability of readability tools to predict L2 reading comprehension is limited and lacks consensus, particularly in exploring the role of cohesion in L2 reading. Additionally, the limited studies available all make use of cloze-testing as a means of reading comprehension, which has been identified as problematic when examining readability tool predictions. The current study aims to fill this gap in the literature by being one of the first studies to use authentic comprehension assessments as indicators of readability tool predictive power. Additionally, the literature discussed thus far only includes the assessment of reading via comprehension tasks given after reading is completed. There are further benefits to be explored through the use of on-line processing measurements to inform the role of cohesion in L2 reading.

**On-line Processing Measurements as Reading Assessment**

Historically, reading is most commonly assessed through off-line comprehension assessments. After the reading has taken place, questions or tasks like free-recall or summarizing are used to draw out the information as understood by the reader. Breakdowns in comprehension therefore are deduced from how the questions are answered. While obtaining off-line data, such as comprehension scores, can give insight into what a reader has processed once the reading has been completed, it does not give insight into the moment by moment decisions that affect the reading process in real time.
Rayner et al. (1989) write, "Understanding how immediate on-line decisions are made will help us to uncover the basic structure and routines of the language processing system" (22). Thus, off-line data are limited in informing how the difficulty of a given text affects the way a reader engages with the text.

In order to obtain insight into the process of reading, a number of on-line tasks have been used in the reading processing literature, including self-paced reading, lexical decision tasks, and fragment completion responses (Clahsen & Felser, 2006). However, each of these tasks involves the reader doing something unnatural to the way silent reading normally occurs, e.g. pressing a button in self-paced reading, providing responses in lexical decision tasks, and interrupting reading with fragment completion responses (Rayner et al., 1989). The last few decades have seen an increase in the number of eye tracking studies as a solution for an unobtrusive means of capturing the real-time processing of silent reading.

In analyzing eye-movement behavior, fixations and saccades are measured to indicate reading. Fixations refer to the amount of time an eye is stationary and fixed on visual input, and saccades are the eye movements between fixations when the eye travels to a new location (Rayner, 2009a). Eye-movements are often used to indicate the cognitive processes behind reading and analysis of the eye-movement record can indicate processing difficulties (Rayner et al., 1989).

Eye-tracking studies in second language acquisition (SLA) contexts have been increasing in recent years in a wide array of research areas. Such SLA studies include investigations in bilingual lexical access (Felser, Sato, & Bertenshaw, 2009); resolution
of syntactic ambiguity (Dussais & Sagarra, 2007; Frenck-Mestre, 2002) and scan path while answering comprehension questions (Bax & Weir, 2012). (For a full review of recent eye-movement studies in SLA research, see Winke, Godfroid, & Gass, 2013). Yet, within SLA eye-movement research, there are no studies such as the one proposed here investigating L2 eye movements at the discourse level of reading, and none that examine cohesion. There have been investigations in cohesion in L2, however, using other on-line processing measures, such as self-paced reading.

Crossley, Yang, & McNamara (2014) used a self-paced, non-cumulative, moving window reading task with L2 readers of English. The authors used simplified texts, a beginner and intermediate level, and verified the simplified texts by running Coh-Metrix analyses. They found that texts simplified for cohesion were read faster and comprehended better than the original, authentic texts; however, they found that learner proficiency was a better indicator of reading time than text difficulty alone. This study is instrumental in that it is one of the first looking at cohesion as a factor in L2 reading using on-line research methods.

2.3 Rationale for Current Study and Research Questions

The literature on reading comprehension is extensive and broad, yet for L2 reading, there exists a less comprehensive understanding of what facilitates readers' understanding beyond L2 lexical and syntactic proficiency. Studies intending to draw out reading abilities lack conclusive answers and often conflate text readability with L2 proficiency by only measuring lexical and grammatical complexity. The L1 literature is in agreement that cohesion facilitates low-knowledge readers and some comprehension
recall for all readers; however, there is little research on the extension of these effects into L2 reading. The facilitating effects of content word overlap found in the L1 literature have been minimally carried out in the L2 context through the use of readability tools in predicting comprehension difficulties in cloze-tests, which have been argued to be less reliable instruments. Experimental design taking advantage of on-line processing has been found to be instrumental in accurately capturing reading comprehension; however, there are only a marginal number of L2 reading comprehension studies involving passage reading and on-line processing measurements. This dissertation therefore seeks to advance the research of cohesion effects by addressing the issues of content word overlap and text readability in L2 reading through two experiments.

Experiment one builds upon the existing literature investigating cohesion as a source of text difficulty through the use of readability tools for predicting L2 reader comprehension. The guiding research question for experiment one is:

RQ 1: Do readability tools analyzing cohesion predict L2 learner comprehension across proficiency levels better than traditional readability formulas, as assessed by non-cloze tests?

This study fills a gap in the literature by using more authentic texts and assessment tasks rather than cloze tasks, which may have biased results towards tools only measuring lexical and grammatical complexity. Additionally, unlike previous studies using homogenous learner groups, this study will use a variety of proficiency levels to investigate if results differ as proficiency of the learner increases.
Experiment two investigates the role of cohesion in L2 comprehension in a more narrowly defined manner through the inclusion of on-line processing measurements and passages carefully manipulated for lexical and syntactic complexity as well as amount of content word overlap. Experiment two is guided by these research questions:

RQ 2: Does content word overlap affect discourse-level text processing by L2 readers?

RQ 3: Do the effects of content word overlap vary based on proficiency of the reader?

RQ 4: Do the effects of content word overlap vary based on difficulty of lexical items in the text?

This study expands the L2 reading literature by being one of a small number of studies to examine L2 reading at the discourse level and one of the few studies that explores the role of content word overlap in L2 reading. Whereas previous studies have explored lexical difficulty effects on L2 reading alone, this study also considers content word overlap and any possible interaction between lexical difficulty and content word overlap. Additionally, this study extends the research on how L2 proficiency impacts reading behavior by using multiple proficiency levels. It also contributes to the body of SLA literature in eye-movement research by being one of the first studies to explore discourse-level reading and breaking new ground in using eye-movement measurements to explain cohesion effects in L2 discourse-level reading.

With the results of both experiments, the dissertation explores the textual factors that impede L2 reading comprehension at multiple stages of learner development that
have not previously been researched and shows how these results can be applied in practical settings
CHAPTER 3

METHODOLOGY

In this chapter, the methodology for the entire study is outlined and explained thoroughly. To provide a comprehensive investigation of the research questions, two experiments were conducted. In section 3.1, the first experiment is described, including the description of the assessment, participants, and intended analysis. In section 3.2, experiment two is detailed, including the participants, creation of the reading instrument, and a detailed explanation of the experimental condition of content word overlap.

3.1 Experiment One

Experiment one broadly examines L2 reading comprehension through the comparison of readability formulas and their predictions on texts used in existing, authentic reading comprehension assessments in a university-level intensive English language program. Specifically, this experiment collected the comprehension scores of English language learners on three different tests of the same proficiency level and analyzed the comprehension scores to see if tests yielded significant differences. Then, the texts used in the assessments were analyzed using two different readability formulas. The analysis compared the predictions of Coh-Metrix (RDL2), which measures cohesion, lexical, and syntactic difficulties, to Flesch-Kincaid Grade Level (FKGL) predictions, which measure only lexical and syntactic difficulties. This experiment is loosely modeled after Crossley et al. (2008) who compared RDL2 with FKGL on a set of cloze-tests and
the reading comprehension of Japanese learners of English. This experiment differs from Crossley et al. (2008) in that the texts used for readability analysis in this study are fewer in number, and comprehension is assessed by means of multiple-choice questions rather than a cloze test.

**Methodology**

**Description of reading assessment.** Experiment one used the scores of an existing reading comprehension assessment in place at a language program at a major southeastern university. The language program provides intensive English courses for English as a foreign language (EFL) students not yet accepted into a U.S. university due to limited language proficiency. The language program has six different proficiency levels, from beginner to high-intermediate (Common European Framework of Reference for Languages [CEFR] levels A1 to B2.2).

The reading comprehension assessment was routinely administered by the program at the end of each nine-week term as a means of determining if students were ready to be promoted to the next level. Tests were administered in class by the teacher and were not timed. The procedure for the test was as follows: (1) reading a level-appropriate expository text of around 300-500 words, (2) composing a summary of the text, and (3) answering objectively-scored questions. Objectively-scored questions included multiple-choice and short answer, and they covered various facets of comprehension such as main idea, supporting details, context clues, pronoun reference, and inferencing. The total test was worth 15 points, where a score of nine (60%) or higher
was considered a passing score by the language program. For an example test, see Appendix A.

According to the policy of the language program, students could fail and repeat the same level of reading class a maximum of three times. As a result, three different reading comprehension tests were created for each level to prevent students from repeating a test. Thus, there were 18 tests available for analysis in experiment one, three tests at each of the six levels of proficiency.

**Description of texts.** The expository texts used in the assessment were selected from a published L2 reading textbook series named *Q Skills Series for Success: Reading and Writing* (Oxford University Press, 2011); the series includes six different textbooks, each of which correspond to the CEFR proficiency standards of the six reading levels of the language program. Three texts from each level of textbook were chosen to be used in the reading comprehension assessments on the basis that each topic was assumed to be equally familiar to as many students as possible. Topics that biased particular majors or fields of expertise were avoided to prevent any topic familiarity effects. See Table 3.1 below for a list of titles and word counts for each test.

As previously stated, assessments were given at the end of each nine-week term and rotated through the three-test cycle (Test set A, Test set B, Test set C) to avoid student repetition. To increase the validity of the analysis, comprehension scores from at least three administrations of each test set were collected. Test dates can be referenced in Table 3.1. There is no relationship between tests within a level other than the intended
proficiency level, nor are there any connections amongst tests in set A as compared with set B; all 18 tests are unique and unrelated.

Table 3.1

*Test titles, word counts, and administration dates for assessments in experiment one*

<table>
<thead>
<tr>
<th>Tests</th>
<th>Text Titles</th>
<th>Word Count</th>
<th>Administration Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>Cristiano Ronaldo</td>
<td>319</td>
<td></td>
</tr>
<tr>
<td>2A</td>
<td>A Tall man in a Small Space</td>
<td>427</td>
<td></td>
</tr>
<tr>
<td>3A</td>
<td>Think Before you Toss</td>
<td>472</td>
<td>July 2015</td>
</tr>
<tr>
<td>4A</td>
<td>How Money Can Make you Happy</td>
<td>520</td>
<td>March 2016</td>
</tr>
<tr>
<td>5A</td>
<td>From Student to Employee</td>
<td>533</td>
<td>October 2016</td>
</tr>
<tr>
<td>6A</td>
<td>Caffeine, by the Can</td>
<td>448</td>
<td></td>
</tr>
<tr>
<td>1B</td>
<td>Why do People Love their Cities?</td>
<td>280</td>
<td></td>
</tr>
<tr>
<td>2B</td>
<td>What is an Adult?</td>
<td>337</td>
<td></td>
</tr>
<tr>
<td>3B</td>
<td>Running a Family Business</td>
<td>341</td>
<td>March 2015</td>
</tr>
<tr>
<td>4B</td>
<td>Fear Factor: Success and Risk</td>
<td>472</td>
<td>October 2015</td>
</tr>
<tr>
<td>5B</td>
<td>Problems with Food Labels</td>
<td>545</td>
<td>May 2016</td>
</tr>
<tr>
<td>6B</td>
<td>The Aral Sea</td>
<td>547</td>
<td></td>
</tr>
<tr>
<td>1C</td>
<td>Good Weather, Bad Weather</td>
<td>262</td>
<td></td>
</tr>
<tr>
<td>2C</td>
<td>Can we Trust our Fears?</td>
<td>349</td>
<td></td>
</tr>
<tr>
<td>3C</td>
<td>Unusual Ideas to Make a Buzz</td>
<td>297</td>
<td>May 2015</td>
</tr>
<tr>
<td>4C</td>
<td>&quot;Out of the Box&quot; Ideas for Greener Cities</td>
<td>375</td>
<td>December 2015</td>
</tr>
<tr>
<td>5C</td>
<td>The Promise of Play</td>
<td>462</td>
<td></td>
</tr>
<tr>
<td>6C</td>
<td>Phototruth or Photofiction</td>
<td>524</td>
<td></td>
</tr>
</tbody>
</table>

**Data Collection.** Since test scores were collected over a span of two years, a total of 1,131 comprehension scores were collected. The distribution of these scores by proficiency level and test set are summarized below in Table 3.2. The participants represented by these comprehension scores were 569 English L2 students with varying L1 backgrounds (378 males, mean age: 20.1). See Table 3.3 for participant summary.

Since these comprehension scores were taken from an educational setting rather than a standardized research design, it is important to note that participants did not take all three
tests in one level. Of the 568 participants, most were promoted to the next level (68%), some left the program (18%), and others failed and thus took more than one test within a level (14%).

Table 3.2

*Total reading comprehension scores collected*

<table>
<thead>
<tr>
<th>Level</th>
<th>Test A</th>
<th>Test B</th>
<th>Test C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>56</td>
<td>57</td>
<td>42</td>
<td>155</td>
</tr>
<tr>
<td>3</td>
<td>64</td>
<td>75</td>
<td>66</td>
<td>205</td>
</tr>
<tr>
<td>4</td>
<td>71</td>
<td>85</td>
<td>86</td>
<td>242</td>
</tr>
<tr>
<td>5</td>
<td>94</td>
<td>99</td>
<td>107</td>
<td>300</td>
</tr>
<tr>
<td>6</td>
<td>76</td>
<td>83</td>
<td>70</td>
<td>229</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>1131</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.3

*Summary of experiment one participant language background*

<table>
<thead>
<tr>
<th>L1</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>221</td>
</tr>
<tr>
<td>Chinese</td>
<td>157</td>
</tr>
<tr>
<td>Japanese</td>
<td>74</td>
</tr>
<tr>
<td>Spanish</td>
<td>62</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>12</td>
</tr>
<tr>
<td>Korean</td>
<td>12</td>
</tr>
<tr>
<td>French</td>
<td>8</td>
</tr>
<tr>
<td>Turkish</td>
<td>6</td>
</tr>
<tr>
<td>Hindi</td>
<td>5</td>
</tr>
<tr>
<td>Portuguese</td>
<td>5</td>
</tr>
<tr>
<td>Thai</td>
<td>4</td>
</tr>
<tr>
<td>Nepali</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>568</strong></td>
</tr>
</tbody>
</table>
Analysis Procedure. First, since comprehension scores were compared across test sets, but different participants took each test, proficiency measures of all participants within levels were compared. The language program has a separate placement exam that is taken every nine weeks by all students (Accuplacer, by College Board), and participants’ reading placement scores were analyzed to ensure a homogeneous proficiency level for each set of comprehension scores within a level.

Reading assessment comprehension scores within levels and across test sets were examined. The initial assumption of the language program was that all three tests within a level should be of comparable difficulty due to the text being taken from the same level of textbook. Additionally, the language program did an item-analysis for each test to ensure that no one question received significantly more incorrect answers. Thus, any comprehension score differences found during the analyses would be the result of text readability.

Finally, the readability predictions from both RDL2 and FKGL were calculated for each of the texts used in the test. As a reminder, RDL2 measures three components of a text presumed to cause difficulty as found in Crossley et al., 2008: (1) CELEX word frequency; (2) syntactic similarity; and (3) content word overlap. See Chapter 2 for a review of these measures. The combined measure of these scores results in a difficulty rating for a given text where a higher score indicates that a text is easier to read. FKGL measure only two components of a text: (1) number of syllables per word and (2) number of words per sentence. These factors combine into a grade-level score predicting readability according to the U.S. grade school system where a higher number indicates a higher grade (e.g., 11.0 indicates a text at an 11th grade reading level).
Readability scores of texts were compared to other texts’ readability scores in the same level and were indicative of the readability tool’s prediction of relative difficulty. This scale of relative difficulty was then compared to actual comprehension scores to determine if either RDL2 or FKGL was more accurate in predicting higher or lower comprehension scores.

3.2 Experiment Two

Experiment two sought to build upon the broad investigation of experiment one, which addressed cohesion indirectly through the use of readability tools. To further continue the investigation of cohesion, the research design of experiment two examined cohesion directly with carefully controlled passages and the incorporation of eye-movement measurements to provide a real-time processing picture of the L2 reader and text difficulty. Experiment two answered the following research questions:

RQ 2: Does content word overlap affect text processing by L2 readers?

RQ 3: Do the effects of content word overlap vary based on proficiency of the reader?

RQ 4: Do the effects of content word overlap vary based on difficulty of lexical items in the text?

Participants

Experiment two involved 98 participants (47 males, mean age: 22.7) recruited from a large southeastern university. All L1 participants were undergraduate students enrolled in the university while L2 participants included both undergraduate and graduate
students. Native languages for the L2 participants include: Chinese, Arabic, Spanish, Portuguese, Japanese, French, Russian, Hindi, Romanian, Serbian, Hungarian, Dutch, and German. L1 participants were recruited through university's psychology participant recruitment website and were given course credit as compensation for participation. L2 participants were recruited both from the university and the intensive English language program and were given gift cards in exchange for participation since no course credit could be given. All participants had normal or corrected vision.

Experimental Design

Vocabulary Size Test. All L1 and L2 participants were given a vocabulary size test (VST) used to identify language proficiency (Nation, 2001). Although the test discriminates vocabulary size for native speakers as well, the L1 scores were only used for the purposes of identifying outlier participants.

The VST is a multiple-choice question test designed to assess the receptive vocabulary size of a participant and is freely available at vocabsize.com. The test prompts participants with 140 decontextualized vocabulary items that range from very frequent (items occurring in the 1,000 most frequently-used word range) to infrequent (items in the 14,000 most frequently-used word range) and forces participants to select one definition for the word among five possibilities. Although the test is not a direct measure of reading skill, receptive decontextualized vocabulary knowledge has been shown to be a critical factor in reading abilities (Nation, 2012).

The test results in a score of an approximated receptive vocabulary size. For instance, a score of 6,000 indicates that the participant has a receptive vocabulary size of
roughly 6,000 word families. Nation (2001) defines ‘word family’ as any words sharing a common stem (e.g. nation, national, nationality, nationalize, etc...) consisting of one word family. The rationale for this distinction is that knowledge of one member of a word family increases the likelihood for receptive knowledge of all words within the word family. For comparison, the estimated vocabulary size needed to succeed in university-level reading is 10,000 word families (Nation, 2012).

In the current experiment, participants took the VST in the lab on a computer after completing the eye-tracking reading. The results were used to divide L2 participants into two groups: high proficiency and low proficiency, where the participants with a vocabulary size of 7,000 or greater were considered high, and those whose VST score was below 7,000 were considered low. The rationale for selecting 7,000 as the cut-off was verified by the literature which indicates that the majority of university students of non-European language background succeed in university with a vocabulary size of 6,000-8,000 (Gungor and Yayli, 2016; Nation, 2012). Thus, any participant with a vocabulary size higher than 7,000 would be of university level and would be considered advanced compared to the students in the intensive language program. This data is summarized in Table 3.4 below.

By having two L2 proficiency groups, it can be seen how learners make use of content word overlap differently based on proficiency, a distinction that was significant for reading studies described in chapter 2 of this manuscript. While the highest proficiency level in experiment one was high-intermediate, experiment two extends the learner proficiency range to advanced levels to see if there are any ceiling effects on cohesion with more proficiency learners.
Table 3.4

Summary of participants in experiment two

<table>
<thead>
<tr>
<th>Group</th>
<th>Vocabulary Size Test Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
</tr>
<tr>
<td>L1</td>
<td>34</td>
</tr>
<tr>
<td>High L2</td>
<td>33</td>
</tr>
<tr>
<td>Low L2</td>
<td>31</td>
</tr>
</tbody>
</table>

The rationale for using native-speaking readers is that it establishes a baseline measurement of successful reading by which to compare the different L2 groups.

**Background Questionnaire.** To collect background information, a background questionnaire was administered to each participant following the VST. The questionnaire is adapted from the Bilingual Language Profile (BLP), and asks participants to self-report language usage, preferences, and language dominance of bilingual adults (Birdsong, Gertken, & Amengual, 2012). The questionnaire also includes language background, additional languages learned, and length of time studying English for nonnative participants.

**Reading Experiment Items.** The passages used in the eye-tracking experiment are 48 five-sentence expository paragraphs between 60-70 words. The justification for using expository texts is that the discourse model in reading expository texts is different than in narrative texts, which rely more heavily on integrating distant information (Coté, Goldman, & Saul, 1998). Passages were created by the researcher and are not the product of modifying existing texts. Passages were written to imitate texts used in recreational
reading rather than learning contexts and are not intended to teach readers or resemble passages from textbooks. Since background knowledge is a critical predictor of reading comprehension, topics selected for experiment two were those found in news headlines, magazine briefs, and other general audience sources; they are varied across interest area and geographical reference and are not designed to create any advantage for readers with certain background knowledge.

**Background Knowledge.** To ensure that background knowledge was not a confounding factor in this analysis, a post-hoc analysis of topic familiarity was conducted to ensure that all passage topics were of equal familiarity to both native and non-native speakers. A random sample of L1 (n=31) and L2 (n=33) participants who did not participate in the eye-tracking experiment took an online survey to self-assess their familiarity of the 48 topics used in the eye-tracking experiment. The survey provided the first sentence of each of the 48 passages, all of which clearly introduce the topic of the passage, and asked participants to respond using a six-point Likert scale to self-rate their familiarity with this topic. Participants were required to choose from the following choices:

1. This topic is very unfamiliar.  
2. This topic is somewhat unfamiliar.  
3. This topic is slightly unfamiliar.  
4. This topic is slightly familiar.  
5. This topic is somewhat familiar.  
6. This topic is very familiar.
Ratings were collected and analyzed by language group (native speakers vs. nonnative speakers). Results showed similar ratings for all 48 passages with no passage receiving a mean score above a four ("This topic is slightly familiar"). Means were similar for both L1 and L2 participants in addition to similar means across the three passage difficulty levels used in the eye-tracking experiment. Thus, it was assumed that topic familiarity for all passages in experiment two was constant and background knowledge did not impact reading measures. Table 3.5 summarizes the ratings by both L1 and L2 participants by passage difficulty. See Appendix B for a complete listing of topics and background ratings for a summary of topics and participant ratings.

Passage structure. Passage metrics were carefully controlled to allow for clear conclusions to be made without possible confounding factors. The structure of each five-sentence passage is identical: a topic-introducing sentence begins each paragraph and is followed by three supporting sentences, each containing a critical word that refers to a common entity. The final sentence contains the target word and is the same across all experimental conditions so that reading times can be directly compared.

Table 3.5

<table>
<thead>
<tr>
<th>Passage Difficulty</th>
<th>Familiarity Ratings L1</th>
<th>Familiarity Ratings L2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Easy Passages</td>
<td>2.19</td>
<td>.72</td>
</tr>
<tr>
<td>Intermediate Passages</td>
<td>2.68</td>
<td>.75</td>
</tr>
<tr>
<td>Advanced Passages</td>
<td>2.98</td>
<td>.68</td>
</tr>
<tr>
<td>All Passages</td>
<td>2.62</td>
<td>.72</td>
</tr>
</tbody>
</table>
All target words are between 4-11 characters long and located near the end of the sentences with a three to four-word spillover region to allow for wrap-up processing (Rayner & Morris. 1990). All targets were nouns and occur within the 2,000 most frequently used words in English in order to have a higher probability of being familiar for all participants. See Figure 3.1 for an example passage.

![Topic Sentence (without critical word)](image)

**Figure 3.1** Example passage illustrating locations of critical and target words.

It is important to note that critical/target words do not appear in the first sentence of any passage. This will disambiguate regressions to the first sentence for the purpose of recalling the general topic of the paragraph from regressions made to refer back to the common referent of the critical/target words.

As the core question of this dissertation asks what features of a text increase difficulty, the following text features were carefully controlled using RDL2 and its sub indices: (1) Cohesion, measured by content word overlap in adjacent sentences in all passages; (2) Syntactic difficulty, measured in Coh-Metrix by sentence syntax similarity of adjacent sentences; and (3) Lexical difficulty, measured by the mean CELEX of content words, excluding critical words. The manipulation of these features described below resulted in three passage difficulty levels: easy, intermediate, and advanced.
**Content word overlap.** Cohesion is addressed in RDL2 and in this study as content word overlap, or the repetition of content words in adjacent sentences. Content word overlap manipulation, as it addresses the research questions, was the primary experimental variable and is further explained in the next section. However, it was also important to ensure that there were no other instances of content word overlap in the passage affecting reading difficulty. Thus, critical/target words were removed from passages, and content word overlap of the remaining text was analyzed to ensure consistency across the three difficulty levels. Measures of content word overlap in passages (without critical words) are given below in Table 3.6.

Table 3.6

*Content word overlap in adjacent sentences by passage proficiency level*

<table>
<thead>
<tr>
<th>Passage Difficulty</th>
<th>Content word overlap in adjacent sentences</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy</td>
<td></td>
<td>0.06</td>
<td>.02</td>
</tr>
<tr>
<td>Intermediate</td>
<td></td>
<td>0.06</td>
<td>.02</td>
</tr>
<tr>
<td>Advanced</td>
<td></td>
<td>0.05</td>
<td>.02</td>
</tr>
</tbody>
</table>

**Syntactic Difficulty.** The primary research question in this study is exploring the interaction between lexical difficulty and cohesion. Thus, for all passages across proficiency levels, syntactic complexity is held constant so as to not be the source of reading difficulty. As in Crossley et al. (2008), syntactic complexity is measured by the similarity of sentence syntax in adjacent sentences. This number is secondarily supported by number of words in each sentence. Measures of syntactic complexity are presented in Table 3.7.
Table 3.7

Syntactic difficulty measurements by passage proficiency level

<table>
<thead>
<tr>
<th>Passage Difficulty</th>
<th>Syntactic Similarity</th>
<th>Words per sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Easy</td>
<td>0.14</td>
<td>.03</td>
</tr>
<tr>
<td>Intermediate</td>
<td>0.15</td>
<td>.03</td>
</tr>
<tr>
<td>Advanced</td>
<td>0.14</td>
<td>.03</td>
</tr>
</tbody>
</table>

**Lexical Difficulty.** Passages were written such that the most meaningful difference between proficiency levels was the mean CELEX frequency of content words found in the passage. Easy passages were written with more frequent words, intermediate passages with less frequent words, and advanced passages used the least frequent words. The mean lexical frequency for each passage difficulty level falls within the proficiency range norms of McNamara, Graesser, McCarthy, and Cai (2014) for corresponding difficulty levels. As was done when measuring content word overlap, critical/target words were removed prior to the analysis since the high frequencies of the critical/target words would have impacted the overall lexical frequency means. The CELEX norms for corresponding difficulty levels and mean lexical frequency of passages without critical words are shown in Table 3.8.

All target words regardless of passage proficiency fall into the easiest CELEX range, 2.398-2.522, so that the reading times collected on these words will be indicative of content word overlap manipulations and not due to lexical knowledge. In sum, passage difficulty across conditions is uniform and is sufficiently distinct to have three well-defined difficulty levels of passages. The above statistics can be supported by the RDL2 measures for all passage difficulty levels.
Table 3.8

**CELEX frequency of content word norms (McNamara et al., 2014) and means by passage proficiency level**

<table>
<thead>
<tr>
<th>Passage Difficulty</th>
<th>CELEX lexical frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Norms</td>
</tr>
<tr>
<td>Easy</td>
<td>2.398-2.522</td>
</tr>
<tr>
<td>Intermediate</td>
<td>2.235-2.339</td>
</tr>
<tr>
<td>Advanced</td>
<td>2.114-2.208</td>
</tr>
</tbody>
</table>

Table 3.9 lists the means for passage RDL2 as well as the norms from McNamara et al. (2014).

Table 3.9

**RDL2 norms (McNamara et al., 2014) and means by passage proficiency level**

<table>
<thead>
<tr>
<th>Difficulty Level</th>
<th>RDL2 measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Norms</td>
</tr>
<tr>
<td>Easy</td>
<td>22-27</td>
</tr>
<tr>
<td>Intermediate</td>
<td>15-19</td>
</tr>
<tr>
<td>Advanced</td>
<td>11-13</td>
</tr>
</tbody>
</table>

**Experimental Conditions.** With syntactic complexity held constant and lexical complexity defining the passage difficulty levels, content word overlap is varied across four different text conditions. The overlap text conditions are:

(N) No Overlap -- in the four sentences following the topic sentence, the same idea is represented by four different nouns (e.g., *plan, method, approach, strategy*);
(C) Close Moderate Overlap -- in the four sentences following the topic sentence, the last two sentences represent the same idea using the same noun while the first two sentences use two different nouns (e.g., *plan, method, strategy, strategy*);

(D) Distant Moderate Overlap-- in the four sentences following the topic sentence, the first and final sentence use the same noun to represent the same idea while the second and third use two different nouns (e.g., *strategy, plan, method, strategy*);

(A) All Overlap -- the four sentences following the topic sentence use the same noun to represent the same idea (e.g., *strategy, strategy, strategy, strategy*).

See figure 3.2 below for an example passage in all four conditions.

The rationale for this design is to ensure that the reading of the final instance of the target word can be attributed to either amount or quality of overlap. This design distinguishes between global (distant) and local (close) cohesion effects as well as effects for amount of cohesion. An effect in the all overlap condition but not the close overlap condition indicates the importance of quantity of overlap; however, an effect in close overlap but not distant overlap indicates the power of local over global cohesion, regardless of quantity.

Each difficulty level (easy, intermediate, advanced) has all four experimental conditions in order to ascertain whether cohesion effects, if present, interact as lexical complexity increases. Passages were randomized across difficulty and condition and separated into a counter-balanced set. Each set was then randomized with each participant. The summary of passages read by participants is summarized in Table 3.10 below. For a list of all passages in all conditions, see Appendix C.
Figure 3.2 One example passage presented in all four overlap conditions. Target word is in red, critical words in green, spillover region in light blue, and topic sentence in dark blue.

**Comprehension Questions.** To assess comprehension and to ensure that participants were truly reading passages, every passage had two comprehension questions. All comprehension questions were true/false and were consistent across experimental conditions. Each passage had one question addressing main idea and one question addressing a detail from the text that was unrelated to the target word. Correct responses were randomized across passages.
Table 3.10

Summary of passage reading by experimental condition and difficulty

<table>
<thead>
<tr>
<th>All Overlap</th>
<th>Close Moderate Overlap</th>
<th>Distant Moderate Overlap</th>
<th>No Overlap</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy (4)</td>
<td>Easy (4)</td>
<td>Easy (4)</td>
<td>Easy (4)</td>
<td></td>
</tr>
<tr>
<td>Intermediate (4)</td>
<td>Intermediate (4)</td>
<td>Intermediate (4)</td>
<td>Intermediate (4)</td>
<td></td>
</tr>
<tr>
<td>Advanced (4)</td>
<td>Advanced (4)</td>
<td>Advanced (4)</td>
<td>Advanced (4)</td>
<td>48</td>
</tr>
</tbody>
</table>

Passage Pilot Study. Since passages were written by the experimenter, a pseudo-pilot study was done to ensure that passages were comprehensible, comprehension questions were valid, and the overlap conditions did not appear unnatural during reading. To accomplish this, passages were read by 87 native speaker participants who were recruited from a first-year English course at the university. None of these participants participated in the actual eye-tracking experiment. Passages were imported into a survey using Qualtrix and were randomized across a counterbalanced set such that each participant only read 48 passages, 12 of each experimental condition. The task was done on a regular computer in a classroom setting and was not done with eye-tracking equipment. Participants read the passage and then rated the passage by responding to the prompt "How easy was it to read this passage?" on a four-point scale with the following responses: 'easy', 'somewhat easy', 'somewhat difficult', 'difficult.' Participants then answered the two comprehension questions for the passage. After reading the 48 passages, participants were asked to respond with what they thought the survey was testing.
All passages regardless of overlap were rated 'easy' or 'somewhat' easy by the L1 participants ($M=2.1$, where 1='easy' and 4='difficult'). From these results, the passages were deemed to be readable as written and that overlap manipulations did not result in awkward or difficult to read passages. Comprehension was stable across questions ($M=95\%$); one passage had one question with a mean accuracy below 80%. This question was subsequently replaced with a new comprehension question. Finally, no participant was able to articulate the purpose of the research study or mentioned noticing the repetition or lack of repetition of words. Thus, no filler passages were added to the experiment.

**Procedure.** All testing for experiment two was done on the university campus in an eye-tracking lab, and all participants were run individually by the experimenter. L1 reader data was collected first from March - May 2017, and L2 data was collected from May - October 2017. Participants first completed the reading task, followed by the vocabulary size test, and lastly the background questionnaire, all in one sitting with breaks as requested. Average amount of time spent on the whole experiment was approximately 45 minutes for native speakers and 1.5 hours for non-native speakers.

**Apparatus.** Eye movements were recorded via an SR Research Eyelink 1000 eye tracker (spatial resolution 0.01°) sampling at 1000 Hz. Participants were seated 90 cm away from a 20-inch monitor with the use of chin rests to minimize head movements. Although viewing was binocular, eye movements were recorded from the right eye only. The task was designed and controlled with SR Research Experiment Builder software. Participants’ eye movements were calibrated followed by validation using a standard nine-point grid at the start of the experiment and after any break.
**Reading task.** At the beginning of the reading task, participants were instructed to read at a normal rate as though reading for pleasure and that they would not be able to go back to the passage after seeing the comprehension question screen. Participants were randomly assigned a counterbalanced set and then read all 48 passages in one randomized block, with a brief two-minute break halfway through and other rest breaks as requested by the participant. Reading times were collected by the software, but there was no time limit placed on the experiment.

Each passage was adjusted to fit on one screen so no button presses were required while reading. When participants were finished with the passage, they fixated on a corner of the screen to move to the questions page. Participants used the keyboard to answer two true/false comprehension questions about the text. Participants were not able to back-track while they answered the questions so as not to conflate reading behavior with scanning for answers to questions. Between each passage was a drift correction screen to ensure accurate eye-movement capture.
CHAPTER 4
RESULTS

In this chapter, the quantitative results are reported from experiment one and experiment two.

4.1 Experiment One Results

Data Analysis

Data screening. Prior to analyzing the comprehension scores from the participants, several analyses were done to determine if data or participants needed to be removed. Level 1, the lowest level of students in the language program, was not included in the analysis due to a low number of test scores available for analysis (N=53). Since beginning learners' comprehension scores are particularly volatile, the lower total number of scores available in comparison to the other levels warranted its removal. Thus, the analysis presented below includes only levels two through six.

Proficiency within levels. Within levels two through six, participants' proficiency scores used by the language program for placement were compared in order to ensure homogeneity of proficiency across tests. Any participant with proficiency scores of three or more standard deviations above or below the proficiency mean for that level was removed from analysis. This resulted in the removal of 42 test scores across all five levels (.03% of the data). The final numbers included in the analysis by level and test are reported below in Table 4.1, along with average proficiency score.
Table 4.1

Number of participants and mean proficiency scores by level and test

<table>
<thead>
<tr>
<th>Lev.</th>
<th>n</th>
<th>M †</th>
<th>SD</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>56</td>
<td>48.4</td>
<td>9.9</td>
<td>57</td>
<td>47.5</td>
<td>11.4</td>
<td>42</td>
<td>49.0</td>
<td>11.2</td>
<td>155</td>
<td>48.21</td>
</tr>
<tr>
<td>3</td>
<td>64</td>
<td>65.4</td>
<td>9.3</td>
<td>75</td>
<td>64.1</td>
<td>10.4</td>
<td>66</td>
<td>65.3</td>
<td>9.7</td>
<td>205</td>
<td>64.85</td>
</tr>
<tr>
<td>4</td>
<td>71</td>
<td>85.2</td>
<td>10.1</td>
<td>85</td>
<td>82.4</td>
<td>9.3</td>
<td>86</td>
<td>83.3</td>
<td>10.8</td>
<td>242</td>
<td>82.34</td>
</tr>
<tr>
<td>5</td>
<td>94</td>
<td>93.5</td>
<td>8.3</td>
<td>99</td>
<td>94.5</td>
<td>9.7</td>
<td>107</td>
<td>95.1</td>
<td>9.5</td>
<td>300</td>
<td>94.48</td>
</tr>
<tr>
<td>6</td>
<td>76</td>
<td>104.7</td>
<td>9.1</td>
<td>83</td>
<td>104.4</td>
<td>9.4</td>
<td>70</td>
<td>102.4</td>
<td>10.6</td>
<td>229</td>
<td>103.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1131</td>
</tr>
</tbody>
</table>

† The maximum score on the proficiency assessment is 120.

Within each level, the number of scores used in the analysis is balanced across the three different tests, and a total of 1,131 test scores are included. Since participants do not take all three tests at each level, proficiency averages are shown to indicate that each test within a level has comparable students taking the test. In other words, one test does not have significantly higher- or lower-proficiency students affecting the comprehension results.

Comprehension data screening. The original assessment was worth 15 points: a written summary worth five points and a 10-question objective assessment section worth 10 points. Being only scored by one rater, summary scores were deemed unreliable without any possibility for inter-rater reliability analysis. Therefore, only the objectively scored portion of the test was used in analysis, resulting in a maximum score of 10 points.
**Comprehension Results.** Next, the comprehension scores on the reading assessments are reported. Comprehension scores on the objective portion of the test are reported for all 18 tests by level, where the maximum score is 10 and scores are presented in percentage form for ease of reference. The average percentage and standard deviations by test and level are reported in Table 4.2 below.

Table 4.2

*Mean comprehension scores (%) by tests and levels*

<table>
<thead>
<tr>
<th>Level</th>
<th>Test A</th>
<th>Test B</th>
<th>Test C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>M</em></td>
<td><em>SD</em></td>
<td><em>M</em></td>
<td><em>SD</em></td>
</tr>
<tr>
<td>2</td>
<td>65.9</td>
<td>22.2</td>
<td>63.3</td>
<td>19.4</td>
</tr>
<tr>
<td>3</td>
<td>58.2</td>
<td>23.6</td>
<td>71.4</td>
<td>23.3</td>
</tr>
<tr>
<td>4</td>
<td>64.7</td>
<td>17.4</td>
<td>64.0</td>
<td>18.1</td>
</tr>
<tr>
<td>5</td>
<td>64.7</td>
<td>21.1</td>
<td>62.8</td>
<td>17.4</td>
</tr>
<tr>
<td>6</td>
<td>71.1</td>
<td>17.8</td>
<td>75.1</td>
<td>19.5</td>
</tr>
</tbody>
</table>

Across all levels, the mean comprehension score is around 60%, indicating the relative difficulty of the test at each level is constant. According to the language program's promotion guidelines, 60% on the entire test, both summary and objective sections, is the minimum passing score to be promoted to the next level.

In order to determine if comprehension on tests within a level varied, a one-way analysis of variance (ANOVA) was conducted by level such that the comprehension scores for each of the three tests were compared to one another. For a reminder of participant numbers per level and test, refer to Table 4.1 The significance threshold was set at .05. Results of the ANOVA are reported in Table 4.3.
Table 4.3

Results of ANOVA for comprehension differences on tests within a level.

<table>
<thead>
<tr>
<th>Level</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>1464.278038</td>
<td>732.139019</td>
<td>1.57</td>
<td>0.2105</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>12973.0625</td>
<td>6486.53125</td>
<td>12.9</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3289.370034</td>
<td>1644.685017</td>
<td>5</td>
<td>0.0074</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>260.5840709</td>
<td>130.2920354</td>
<td>0.36</td>
<td>0.7014</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>9533.52749</td>
<td>4766.763745</td>
<td>13.15</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

In levels two and five, there are no significant differences in comprehension score across the three tests. In level two, comprehension scores on all three tests remain constant as in level five. This consistency in score supports the assumption that students comprehended the material in all three of the texts approximately equally.

Within levels three, four, and six, however, comprehension scores differed significantly on at least one of the tests. Post-hoc analyses using Tukey-Kramer comparison found mean comprehension scores for test 3B were significantly higher than 3A ($p < .0001$) and 3C ($p = .0009$) while tests 3A and 3C did not differ from one another ($p = .573$). At level four, test 4C yielded significantly lower comprehension scores than 4A ($p = .017$) and 4B ($p = .018$), while 4A and 4B had no significant differences ($p = .959$). For level six, test 6C also has significantly lower scores than tests 6A ($p = .002$) and 6B ($p < .001$), but 6A and 6B did not differ from each other ($p = .366$).

Thus, the analyses show that comprehension among all of the tests within levels three, four, and six is not equal. Since the test developers had run item analyses on each
test to ensure test question validity, the source of difficulty is not due to more or less difficult tests or test questions. Also, proficiency of participants has been controlled such that even though the same participants are not taking all three tests, the same level of participant is. Therefore, a source of comprehension difficulty worthy to investigate is the readability of the texts.

Readability Results. In order to investigate readability as the source of incongruity in comprehension performance within levels, readability predictions from both Coh-Metrix (RDL2) and Flesh-Kincaid Grade Level Index (FKGL) are reported by level and by test. As a reminder, RDL2 considers content word overlap, mean lexical frequency, and syntactic similarity factored in a composite score that ranges in this analysis from 28.9 (easiest text in this set of texts) to 5.0 (most difficult text in this set of texts). FKGL factors only lexical and syntactic components of the text through number of syllables per word and number of words per sentence for a range in this analysis from 5.1 (easiest text in this set of texts) to 11.9 (most difficult text in this set of tests). Scores for each test are presented in Table 4.4 below.

Overall level means of text readability from both RDL2 and FKGL indicate increasing amounts of difficulty by level. Unsurprisingly, level six tests are the most difficult as predicted by both tools and level two tests are the easiest. FKGL scores increase in a step-wise pattern to indicate more difficult texts at each successive level. RDL2 scores decrease to also show increasing difficulty by level as well, with the exception of levels four and five. Standard deviation is included to show the amount of variability in difficulty estimates from both tools within a level. As a reminder, the scale for FKGL is much narrower than RDL2 and thus results in smaller standard deviations.
Table 4.4

Readability scores for each test by Coh-Metrix (RDL2) and Flesch-Kincaid (FKGL)

<table>
<thead>
<tr>
<th>Level</th>
<th>Test A</th>
<th>Test B</th>
<th>Test C</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2</td>
<td>RDL2 25.312</td>
<td>Test B 28.933</td>
<td>Test C 27.476</td>
<td>27.24</td>
<td>1.822</td>
</tr>
<tr>
<td></td>
<td>FKGL 6.563</td>
<td>5.114</td>
<td>6.212</td>
<td>5.963</td>
<td>0.756</td>
</tr>
<tr>
<td></td>
<td>FKGL 6.188</td>
<td>7.841</td>
<td>8.619</td>
<td>7.549</td>
<td>1.241</td>
</tr>
<tr>
<td></td>
<td>FKGL 9.176</td>
<td>7.169</td>
<td>8.043</td>
<td>8.129</td>
<td>1.006</td>
</tr>
<tr>
<td>Level 5</td>
<td>RDL2 16.158</td>
<td>Test B 17.529</td>
<td>Test C 21.325</td>
<td>18.337</td>
<td>2.677</td>
</tr>
<tr>
<td>Level 6</td>
<td>RDL2 12.959</td>
<td>Test B 10.862</td>
<td>Test C 5.06</td>
<td>9.627</td>
<td>4.092</td>
</tr>
</tbody>
</table>

Of central interest to this analysis is how each tool predicts difficulty within levels. At the time of writing, there are no set standardized scores for either FKGL or RDL2 aligned with CEFR levels; thus, the only analysis that can be conducted is from the current data scores. However, with only three tests for comparison, inferential correlations are impossible to derive. Therefore, this dissertation seeks to carefully analyze the raw means and variance descriptively and make suggestions for future research. This analysis focuses on the relationship between level readability predictions by measuring the distance between predictions. Greater distances between predictions indicate greater ease or difficulty of the text and would predict comprehension score differences were the readability tool to be reliable.

The distance between RDL2 scores of pairs of tests is presented below in Table 4.5. As a reminder, lower RDL2 scores indicate more difficult texts. Thus, a negative
distance indicates RDL2 to predict that test to be a more difficult text; a positive distance indicates RDL2 prediction of an easier text. To simplify the comparison, the significance values from comprehension score differences found in the previous analysis are repeated here for each pair of tests; tests found to be significantly easier or harder are highlighted in the table below for ease of reference. For RDL2 to be considered reliable, larger RDL2 distances between tests should be paired with significant comprehension performance differences. Scores highlighted in gray indicate significantly different comprehension scores for the pair of tests from the analysis of Table 4.2.

Table 4.5
Test-pair differences in RDL2 readability predictions alongside significance values for actual comprehension score differences.

<table>
<thead>
<tr>
<th>Level</th>
<th>RDL2 Difference</th>
<th>Test A - Test C</th>
<th>Test B - Test A</th>
<th>Test C - Test B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2</td>
<td>-2.164</td>
<td>3.621</td>
<td>-1.457</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comprehension score p-value</td>
<td>$p = .185$</td>
<td>$p = .803$</td>
<td>$p = .185$</td>
</tr>
<tr>
<td>Level 3</td>
<td>1.101</td>
<td>7.524</td>
<td>-8.625</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comprehension score p-value</td>
<td>$p = .573$</td>
<td>$p &lt; .001$</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>Level 4</td>
<td>5.261</td>
<td>3.451</td>
<td>-8.712</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comprehension score p-value</td>
<td>$p = .018$</td>
<td>$p = .952$</td>
<td>$p = .018$</td>
</tr>
<tr>
<td>Level 5</td>
<td>-3.167</td>
<td>2.371</td>
<td>0.796</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comprehension score p-value</td>
<td>$p = .945$</td>
<td>$p = .694$</td>
<td>$p = .852$</td>
</tr>
<tr>
<td>Level 6</td>
<td>7.899</td>
<td>-2.097</td>
<td>-5.802</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comprehension score p-value</td>
<td>$p &lt; .001$</td>
<td>$p = .366$</td>
<td>$p &lt; .001$</td>
</tr>
</tbody>
</table>

An example reading of this table is provided for clarity. In level two, test 2A was 2.16 points lower in RDL2 readability than test 2C, and there were no significant differences between test 2A comprehension scores ($p = .185$). Moving to level three, the RDL2
readability of test 3B was 7.524 points higher than test 3A and 8.625 points higher than test 3C, indicating that 3B was easier than both 3A and 3C. In the comprehension analysis, test 3B had significantly higher scores from test 3A ($p < .0001$) and 3C ($p < .0001$).

For all RDL2 scores presented above, any prediction greater than five points of difference is connected to a test that had a statistically higher or lower comprehension score in the same direction, meaning if RDL2 predicted a test to be more difficult, the test received lower comprehension scores, and a test predicted to be easier received higher comprehension scores. Test 4C is 8.7 points lower than 4B and 5.26 points lower than 4A; comprehension scores mirror these results with 4C having lower comprehension scores (Mean = 56.7) than 4A (Mean=64.7; $p = .018$) and 4B (Mean=64.0; $p = .018$). Test 6C is 5.8 points lower in RDL2 prediction than 6B and 7.9 points lower than 6A. Again, comprehension scores of 6C (Mean = 59.9) are lower than 6B (Mean = 75.1; $p < .001$) and 6A (Mean = 71.1; $p < .001$).

Contrastingly, any RDL2 score differences less than four points did not result in significantly higher or lower comprehension scores. Test 4B is just 3.4 points higher than test 4A, and there were no meaningful differences found when comparing the comprehension scores (4B M=64.0; 4A M= 64.7; $p = .952$). As seen earlier within level five, there were no significant differences found in any comprehension scores, and when looking at RDL2 predictions, the largest gap is between test 5A and 5C of only 3.2 points. At level six, test 6B is 2.1 points lower than test 6A and the comprehension scores reveal no differences (6B M = 75.1; 6A M = 7.11; $p = .366$).
Based on the comparisons illustrated in Table 4.5, the pattern reveals that variation of at least five points in RDL2 score is associated to a test that has meaningful performance differences in the direction predicted by RDL2. However, if the RDL2 scores differ by less than four points, the associated comprehension scores do not differ significantly. If the pattern holds true in further analysis, it would imply that slight deviations in readability as measured by RDL2 do not impede comprehension, but larger gaps in prediction reliably predict comprehension performance. This is only an observation based on the data trends and requires more than three tests at each level for statistical confirmation.

The data thus far tentatively support the RDL2 readability predictions, and now FKGL predictions are analyzed in the same way. As stated earlier, differences in FKGL scores will be smaller due to the smaller range in FKGL's scale. Also, FKGL scores increase with difficulty; thus, any negative differences shown here indicate the relative ease of test and positive differences indicate more difficulty. The differences in FKGL scores between tests are found in Table 4.6 along with the same p-values on the comprehension score differences. Highlighted cells indicate the tests where comprehension differences were found; bold cells indicate greatest differences between FKGL predictions.

When looking at FKGL predictions, we do not see similar trends as with RDL2. The largest distances in FKGL predictions are for test-pairs: 3C and 3A (2.43), 4B and 4C (2.007), 5A and 5C (-2.341), as well as 6B and 6A (2.45). Yet, none of these test pairs showed significant comprehension differences.
The only large FKGL distance that has significant comprehension score differences is between 6B and 6C of -2.1, where test 6B is accurately predicted to have higher comprehension. However, FKGL also predicts 6B to have higher comprehension than 6A (a distance of 2.451) although there is no difference found in comprehension scores ($p = .366$). In sum, FKGL readability predictions do not consistently predict the tests where participants perform differently as RDL2 does.

Inferentially demonstrating these results was not possible due to only having three tests per level. Future testing should allow for additional tests per level to inferentially assert these claims through a correlational assessment. However, descriptively we can see that when larger variations in RDL2 scores exist, L2 reading comprehension follows the direction of the score whereas FKGL scores do not.
4.2 Experiment Two Results

Experiment two was designed to explore the components of RDL2 in reading comprehension by examining how content word overlap affects reading in passages of careful design. Unlike experiment one, passage metrics affecting RDL2, syntactic similarity, lexical frequency, and content word overlap, have been controlled so as to draw conclusions. The research questions guiding experiment two are presented again below:

RQ 2: Does content word overlap affect discourse-level text processing by L2 readers?

RQ 3: Do the effects of content word overlap vary based on proficiency of the reader?

RQ 4: Do the effects of content word overlap vary based on difficulty of lexical items in the text?

In this section, the procedures used to ensure data reliability will be explained. Next, the results of the overall reading measures will be presented to establish group and passage difficulty differences, and then finally the on-line measures localized to target word reading time will be presented.

Analysis

**Data Screening.** In order to have the most reliable data, several procedures were followed for removing participants and data points. Participants were eliminated if total reading time was less than three standard deviations from their proficiency group mean as
this would be evidence of not reading. Additionally, any participant's total comprehension score that was more than three standard deviations below the mean was removed as evidence of not reading during the experiment. These procedures resulted in the removal of two low L2 participants. Additionally, one L1 participant had a vocabulary size score more than three standard deviations below the group mean (VST=6300) and had total reading times more than four times above the L1 mean. Thus, this participant was removed from analysis.

Certain data points were also removed from analysis. As per standard eye-tracking experiments, any fixation less than 120 ms was removed. Standard procedure in eye-movement research recommends any fixation greater than 800 ms to be removed prior to analysis; however, this is standardized for L1 reading only. In this analysis, fixations greater than 3000 ms were removed prior to analysis. For gaze duration, any measurement more than three standard deviations from the proficiency group mean was removed, resulting in the removal of .01% of the data. In regression path duration, any fixation more than three standard deviations from the mean was also removed, resulting in .2% of the data being removed.

All fixations in gaze duration and regression path duration were only included if the target word was fixated prior to a fixation on any word to the right of the target as recommended in Rayner (2009). As such, all measurements reflect the first encounter of the target and are not the product of rereading. Any instances where the target word was skipped entirely or skipped initially and then revisited are treated as missing data and not included in the analysis. This resulted in the removal of 3% of the data. A second pass check was done to ensure that there remained at least three fixations per condition per
participant so as to have a stable representation of the measurement, so a final 2% of the remaining data were removed due to a single participant having fewer than three fixations in one condition. The final participant numbers are L1 n=34; High L2 n=33; Low L2 n=29.

Global Reading Measures. First, global reading measures were analyzed to determine if participant groups were distinct from one another and if passage difficulty levels were distinct. To determine this, three overall reading measures were used: fixations per passage, total reading time, and comprehension. Results are presented by participant group and passage difficulty in table 4.7 below.

To establish whether global reading measures differed by group and by passage difficulty, repeated measures mixed effects analyses of variance (ANOVAs) were conducted to find the effects of group and passage difficulty on fixations per passage, total reading time, and comprehension. All effects were statistically significant at the .05 significance level. Main effects were revealed for both group and passage difficulty in all three dependent measures.

Fixations per passage refers to the total number of eye fixations made throughout the text, where increased number of fixations is indication of more processing difficulties (Rayner, 2009). There was a main effect for group within fixations per passage \[ F(2,95)=894.29, p < .0001 \].
Table 4.7

*Overall reading measures by group and passage difficulty*

<table>
<thead>
<tr>
<th></th>
<th>Easy Passages</th>
<th>Intermediate Passages</th>
<th>Advanced Passages</th>
<th>Total Avgs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixations per Passage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>L1</td>
<td>60.6</td>
<td>19.6</td>
<td>69.0</td>
<td>21.8</td>
</tr>
<tr>
<td>High L2</td>
<td>81.5</td>
<td>28.8</td>
<td>95.9</td>
<td>37.8</td>
</tr>
<tr>
<td>Low L2</td>
<td>111.4</td>
<td>44.0</td>
<td>129.4</td>
<td>45.2</td>
</tr>
<tr>
<td>Total Reading Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>L1</td>
<td>12.6</td>
<td>4.5</td>
<td>14.7</td>
<td>5.0</td>
</tr>
<tr>
<td>High L2</td>
<td>18.8</td>
<td>7.7</td>
<td>22.6</td>
<td>10.1</td>
</tr>
<tr>
<td>Low L2</td>
<td>28.7</td>
<td>12.2</td>
<td>34.1</td>
<td>13.4</td>
</tr>
<tr>
<td>Comprehension (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>L1</td>
<td>96.6</td>
<td>13.3</td>
<td>94.4</td>
<td>17.3</td>
</tr>
<tr>
<td>High L2</td>
<td>94.1</td>
<td>16.8</td>
<td>91.0</td>
<td>21.2</td>
</tr>
<tr>
<td>Low L2</td>
<td>83.6</td>
<td>27.5</td>
<td>68.3</td>
<td>33.9</td>
</tr>
</tbody>
</table>
A post-hoc analysis using the Tukey-Kramer comparison indicated that native speakers made fewer fixations per passage than the high L2 group \((p < .0001)\) and the high L2 group made fewer fixations than the low L2 group \((p < .0001)\). There was also a main effect for passage difficulty \([F(2,194)=83.07, \ p < .0001]\) where fewer fixations were made in easy passages than both intermediate \((p < .0001)\) and advanced passages \((p < .0001)\). However, there were no significant differences in number of fixations between intermediate and advanced passages \((p = .09)\).

Total reading time measured the amount of time spent reading the passage and does not include time spent answering questions. There was a main effect for group \([F(2,95)=1195.9, \ p < .0001]\), where native speakers read faster than high L2 readers \((p < .0001)\) who in turn read faster than low L2 readers \((p < .0001)\). For passage difficulty, a main effect was also found \([F(2,194)=79.16, \ p < .0001]\). Easy passages were read faster than both intermediate \((p < .0001)\) and advanced passages \((p < .0001)\); however, no meaningful differences in reading time were found between intermediate and advanced passages \((p = .16)\).

For comprehension score, a chi-square analysis was conducted and a main effect for group was found, \(X^2(2, N=95)=47.56, \ p < .0001\). No meaningful difference between high L2 comprehension and L1 comprehension was found \((p > .05)\); however, low L2 readers comprehended significantly less than both groups \((p < .0001)\). The main effect of passage difficulty was also significant \(X^2(2, N=95)=42.06, \ p < .0001\). Easy passages were comprehended with greater success than intermediate or advanced passages \((p < .0001)\), but no difference was found for comprehension between intermediate and advanced passages \((p > .05)\).
In sum, these results indicate that the three proficiency groups are distinct groups that pattern differently from each other in global reading measures. It further supports the proficiency distinction between low and high L2 groups, where high L2 participants make fewer fixations, read faster, and comprehend more than their lower proficiency counterparts. Interestingly, although high L2 readers made more fixations and read more slowly, they comprehended as well as the L1 group.

Additionally, these results indicate passage difficulty distinctions. Easy passages were read faster with fewer fixations and comprehended better than intermediate or advanced passages. However, there are no significant differences in passage fixations, reading time, or comprehension between intermediate and advanced passages, suggesting that the lexical frequency criteria is not distinct enough between the two passage levels to result in meaningful differences. With these understandings of group and passage differences, we now turn to the target reading times and the experimental conditions.

**Local reading measures: Target word results.** In this section, on-line processing measurements on the reading of target words is examined. Two main analyses are examined: gaze duration, which is the sum of all fixations made on the target word prior to exiting the word and is indicative of early reading processing, and regression path duration, which is the summed total time spent from first entering the target word to any regressive fixations earlier in the text prior to moving to a word after the target and is usually used in the eye-movement literature to indicate integration of the target into the context. These measures are discussed in light of group, passage difficulty, and experimental overlap differences. As a reminder, target words are consistently all 5-9 characters across all passage difficulties and located in the same position across all
experimental conditions. Below are examples of participant data to exemplify both measures being used. An example of gaze duration is seen in Figure 4.1 and regression path duration in Figure 4.2

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{gaze_duration.png}
\caption{
Example of gaze duration
\end{figure}

Red fixations were made during first pass reading and are included in gaze duration measurement. The purple fixation was made on a second reading of the target and is not included in gaze duration.

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{regression_path.png}
\caption{
Example of regression path duration
\end{figure}

Purple fixations are first-pass reading fixations, green fixations are second-pass reading, and blue fixations are third-pass reading. At the target word 'status,' the two purple fixations triggered a rereading of the line (green fixations) before advancing to the word to the right of the target 'in.' Thus, regression path duration is the sum of the purple fixations on 'status' plus all green fixations.

**Gaze duration.** First, mean gaze duration times are presented by group and passage difficulty. Mean reading times are shown in Table 4.8.
Table 4.8

Mean gaze durations (msec) on target words by passage difficulty

<table>
<thead>
<tr>
<th></th>
<th>Easy</th>
<th>Intermediate</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passages</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>L1</td>
<td>226</td>
<td>93</td>
<td>242</td>
</tr>
<tr>
<td>High L2</td>
<td>260</td>
<td>115</td>
<td>297</td>
</tr>
<tr>
<td>Low L2</td>
<td>346</td>
<td>150</td>
<td>403</td>
</tr>
</tbody>
</table>

Results of a mixed ANOVA indicate a main effect for both group \(F(2,95)=390.27, p <.0001\) and passage difficulty \(F(2,194)=29.88 p <.0001\). Native speakers made shorter fixations on the target word than the high L2 readers \(p<.0001\), who made shorter fixations than the low L2 readers \(p <.0001\). For passage difficulty, target words found in easy passages had shorter fixations than intermediate \(p <.0001\) or advanced passages \(p <.0001\) but there were no meaningful differences between intermediate and advanced \(p=.248\). These findings confirm gaze duration as a meaningful measure. All target words, regardless of passage difficulty, are the same length and of the highest frequency; in isolation they would be processed approximately the same. However, in the more or less difficult contexts based on lexical difficulty, processing times vary. Thus, the longer fixation times on targets in the intermediate and advanced conditions reflects the greater processing burden of the more difficult passage level.

Next, gaze duration measurements are explored through the experimental variable of overlap. For ease of referencing, one passage in all four conditions is shown below in
Figures 4.3 - 4.6 with example participant eye fixations; one passage is an L1 participant with (Figure 4.3), two are high L2 participants (Figures 4.4 and 4.5), and one is a low L2 participant (Figure 4.6). The blue circles indicate fixations with size relative to duration of fixation. The yellow lines indicate saccades. The arrows at the ends of saccades indicate the direction of saccade and whether it was a forward or regressive movement.

In the figures, note the number of fixations and regressive saccades for each proficiency level of learner. Also note the location of the target word in each example and how the critical words shift based on overlap condition.

*Figure 4.3.* L1 participant example passage with eye-movements, no overlap
Discount airlines, like Ryanair, are finding new ways to make money. The airlines’ strategy is to charge people more money for other benefits. Their strategy has been to have the lowest prices on flights. After people buy the ticket, the airlines’ strategy has been to have additional fees for luggage. The companies have made a lot of money with their strategy in the last year.

**Figure 4.4.** High L2 participant example passage with eye-movements, all overlap

**Figure 4.5** Low L2 participant example passage with eye-movements, distant overlap
In terms of how overlap influenced gaze duration on target words, we find a main effect [$F(3, 285) = 4.19 \ p = .006$]. Separate within group ANOVAs find that these significant differences occur in the high L2 group ($p=.04$) and low L2 group ($p = .026$) but not for the L1 group ($p=.59$). In the low L2 group, there is an incremental consistent increase in fixation time as overlap decreases in both quantity and quality; however, the only significantly meaningful difference is that target words in no overlap were fixated for longer than target words in the all overlap condition ($p=.023$). For high L2 readers, the all, close, and distant conditions have no significant difference in gaze duration time but target words in the no overlap condition are read significantly slower than those in the close condition ($p =.016$). As stated before, there were no significant differences in target reading time for the L1 group; however, the descriptive means indicate that target words
in the no overlap condition were read the fastest. Mean gaze duration times are depicted below in Figure 4.7.

![Figure 4.7 Mean gaze duration times by overlap and proficiency group.](image)

There were no interaction effects for passage difficulty and gaze duration, and this is likely due to the great amounts of variance present in the data, particularly in the low L2 data. However, a descriptive summary of the results will be presented here as the overlap patterns found across all passages shift when lexical difficulty is considered.

For the low L2 group, the effects of overlap on gaze duration are different when looking at easy passages. Target words in easy passages are read more quickly when occurring in positions of all overlap or close overlap. Thus, when there is little lexical complexity, low L2 readers benefit equally from four instance of overlap as well as two instances, provided the overlap occurs in adjacent sentences. This is different from the
main trend found across all passages, where targets in all passages were consistently read faster than close.

However, as lexical difficulty increases in the intermediate and advanced passages, there are clear differences between all overlap, moderate overlap (close and distant overlap), and no overlap. When lexical complexity increases, this benefit of moderate adjacent overlap found in easy passages disappears, and only targets with four instances of overlap result in faster reading time. Moderate overlap still retains a benefit over no overlap, but the reading time differences between close and distant conditions is diminished, implying that moderate overlap, regardless of its location, is moderately beneficial in lexically complex passages.

Additionally, significant differences that were found when comparing easy to intermediate passages are no longer present when taking overlap into account. When intermediate or advanced passages contain all overlap, reading times of target words are not significantly different than targets found in easy passages ($p > .05$). Thus, the presence of increased overlap interacts with increased lexical difficulty to result in faster target reading time. The interaction of overlap and lexical difficulty for low L2 readers is summarized in Figure 4.8.

**Regression Path Duration.** In addition to gaze duration, regression path duration measurements were also collected, which indicate how long a learner spent looking back in the passage after encountering the target. The results of the ANOVA indicate there was a main effect of group on regression path duration [$F (2, 95) = 52.34 \, p < .0001$].
Figure 4.8  Low L2 target reading time (msec) by passage overlap and difficulty

Between groups, low L2 learners spent the most time regressing ($p < .0001$) while regression duration was comparable between the high L2 and L1 groups with no significant differences ($p > .05$). As in previous measures, there was a main effect for passage difficulty from all three groups on duration of regressions [$F(2,194) = 4.96, p = .0008$], where the more difficult passages resulted in more regressive reading from the target. This data is summarized in Figure 4.9.

Next, regression path duration measures are examined in the experimental overlap conditions. There was no main effect of overlap on regression duration [$F(2,285) = .59, p=.622$]. Within groups, there were no significant differences for the low L2 group between the all, distant, and close conditions while the no overlap condition resulted in significantly more rereading ($p = .045$). There were no significant differences in condition for the high L2 and L1 group, although the no overlap condition resulted in the
longest regression reading times for the high L2 group and the shortest regression reading times for the L1 group. These patterns are represented below in Figure 4.10.

![Figure 4.9](image1.png)  
*Figure 4.9* Regression path duration from the target word by group and difficulty

![Figure 4.10](image2.png)  
*Figure 4.10* Regression path duration from the target word by group and overlap
When looking at regression durations from the interaction of overlap and difficulty, there is no main effect \(F(6, 569) = .64, p=.699\). However, interesting patterns emerge. In the low L2 group, intermediate regression path duration averaged 490ms while easy passages averaged 446ms. The all overlap in intermediate passages, however, had a mean regression path duration time of 463ms, an insignificant difference from easy passages. Thus, the presence of all overlap facilitated the integration of the target in lexically complex passages to a processing time comparable to a lexically simpler passage.

In sum, the on-line processing measures of gaze duration and regression path duration revealed significant differences for amount of overlap present, namely that less overlap affects low L2 readers in both target gaze duration time and regression path duration from the target, while only the total absence of overlap affects high L2 readers. Additionally, while the lexical complexity of passages affects gaze duration and regression path duration of all three groups, these effects are mitigated in nonnative speakers when analyzing the interaction of cohesion and lexical difficulty. These results again indicate that more lexical complexity in passages does not uniformly affect L2 reading behavior and measures of cohesion must also be assessed.
CHAPTER 5

DISCUSSION

This section summarizes the results from both experiments. The first research question is discussed in light of the results of experiment one followed by research questions two through four and the results of experiment two. Then a general discussion ensues, shedding light on the role of content word overlap in L2 reading comprehension by exploring the connection of the results from both experiments. In section 5.4, limitations of this research are considered with suggestions for future directions of research. Lastly, practical applications of this research are made and conclusions are given.

5.1 Experiment One Discussion

Research Question 1 asked:

(1) Do readability tools analyzing cohesion predict L2 learner comprehension across proficiency levels better than traditional readability formulas, as assessed by non-cloze tests?

To address this question, experiment one looked at the data of an authentic reading comprehension assessment made up of three unique tests per proficiency level. Within proficiency levels three, four, and six, statistically significant differences were found in the mean comprehension scores of individual tests, indicating at least one of the
tests was significantly more or less difficult. In levels two and five, no significant
differences were found in test comprehension, indicating all three tests within a level
were comprehended approximately equally. Since the questions used in the assessment
had been assessed for validity, and proficiency ability of students had been confirmed to
be constant across levels and testing administrations, we turn to the readability of the
passage used in the test as explanation for the presence or absence of meaningful
differences. Readability predictions from two different tools were found for each of the
three tests within a level, and distance between the predictions was measured. A
descriptive analysis found that when compared to the comprehension score differences, a
pattern emerges that aligns distance in Coh-Metrix (RDL2) score to performance
difference. Contrastively, the distance between FKGL predictions do not align with
comprehension score differences.

Experiment one results found that when RDL2 predicted a difference of five
points or more between two tests of the same level, comprehension scores were found to
be significantly easier or harder in alignment with the RDL2 prediction (all tests had \( p \)-
values < .05). Any pair of tests within a level with less than four points of difference in
RDL2 readability predictions were not found to be significantly different in
comprehension scores. The pattern in these results supports the claim that RDL2
readability predictions align with actual L2 reading comprehension performance. The
RDL2 scores by level are graphed below in figure 5.1; tests with significant
comprehension differences are noted ( * ). As can be visually seen, RDL2 predictions of
greater distances than their test counterparts are the only tests which results in
comprehension differences. Tests of comparable difficulty as predicted by RDL2 (less than four points) do not result in any comprehension differences.

![Figure 5.1 Graph of RDL2 predictions by level; tests with significant comprehension differences noted with *](image)

An analysis of FKGL predictions on the same tests does not result in such trends. Out of the three tests found with significant comprehension differences (3B, 4C, and 6C), only test 6C was found to have a relatively large gap in FKGL prediction than test 6B (2.14 points); however, test 6C also had significantly lower comprehension scores than test 6A, and this difference was not predicted by FKGL (.31 points). All of the other largest gaps in FKGL prediction (between 2.0 and 2.5 points) were associated to pairs of tests that did not yield comprehension differences. Therefore, these results do not support the notion that FKGL is predictive of L2 reading comprehension. FKGL scores for each test by level are depicted below in Figure 5.2. As in Figure 5.1, tests which resulted in significant comprehension differences are marked by *. It can be seen in the graph that...
tests which resulted in significant comprehension differences are not predicted by the FKGL scores. Also, the largest variability in FKGL scores do not align with tests that were significantly different from one another.

![Graph of FKGL predictions by level; tests with significant comprehension differences noted with *](image)

*Figure 5.2 Graph of FKGL predictions by level; tests with significant comprehension differences noted with *

Experiment one also highlights that small differences in readability tool predictions, four points or less in RDL2 for this study, do not result in significant comprehension differences. It was only when RDL2 predictions were five points or greater that significant comprehension differences were demonstrated. Thus, ranges in readability are to be expected for levels of learner rather than a direct correlation of readability to performance.

The data examined here tentatively support research question one by indicating that RDL2 is more reliable in predicting comprehension differences than FKGL based on descriptive observations. This claim could be solidified in the future by having additional
tests per level such that correlational analyses can be done. These results align with Crossley et al. (2008), who also directly compared RDL2 and FKGL on second language reading comprehension, in that they also found RDL2 to be more reliable than FKGL. They are also supported by the findings of Crossley et al. (2011) in that RDL2 is aligned with the principles used by practitioners utilizing intuitive text simplification whereas FGKL is less reliable.

However, these results depart from previous studies examining readability tools in L2 reading by not finding FKGL a reliable predictor of comprehension differences. Crossley et al. (2008) found RDL2 to be more reliable than FKGL; however, their conclusion was that FKGL was still a reliable predictor as did Greenfield (2004) and Hamsik (1984). However, as previously mentioned, the difference in the present study is the type of assessment tool used in measuring comprehension. Earlier studies supporting FKGL used cloze tests where the current study uses a traditional question-based comprehension assessment. Others have argued that cloze tests may be better aligned with a tool measuring lexical/syntactic features due to the design of cloze tests (Crossley et al., 2007), and the findings of this study align with this claim. The comprehension scores of a traditional reading assessment which does not bias lexical or syntactic processing but instead assesses overall reading comprehension through multiple-choice questions does not appear to align with FKGL predictions. This type of reading assessment more closely resembles real-world reading and is likely to be informative of reading difficulties. These findings also support the closer review of materials used in earlier empirical L2 reading studies which only reported the lexical and grammatical measures as evidence of text difficulty. Including analyses using Coh-Metrix on these
materials may better inform the conclusion of earlier studies on sources of reading difficulty.

What differentiates RDL2 and FKGL is the addition of cohesion in the analysis of difficulty. By factoring cohesion together with lexical and syntactic measures, RDL2 is a more robust predictor of the readability of a text. The results of this study are in line with the reading models indicating the importance of referential cohesion in text integration and overall processing (Kinstch & Van Dijk, 1978). As learners process the lexical and syntactic complexities, content word overlap eases the processing burden by facilitating the integration of information into the discourse model. Rather than needing to make inferences or bridging connections in the discourse on their own, readers are facilitated by the repetition of words throughout the text, allowing easier integration and supporting greater comprehension. Other empirical work has shown that content word overlap has impacted L1 reading for low-knowledge readers, but this study is one of the first to demonstrate its effects in L2 reading (McNamara et al., 2010).

The results of experiment one do not diminish the effects that lexical and syntactic complexity have on reading comprehension, but instead argue that a formula which factors content word overlap in addition to lexical and syntactic complexity is more effective than one which ignores cohesive factors. It should also be noted that these results do not isolate content word overlap as the source of the comprehension score differences, but rather looks at the combined impact when content word overlap interacts with lexical frequency and syntactic difficulty. These three measures are significant as they pertain to the psycholinguistic model of reading through word recognition, sentence
parsing, and integrating information into the discourse model. Experiment one provides evidence that all three should be considered when determining text difficulty.

5.2 Experiment Two Discussion

Experiment two was designed to answer the remaining three research questions using eye-tracking measurements. Results of experiment two will be discussed in light of each of the research questions, beginning with research question two.

Research Question Two Discussion

Research question two is restated here for ease of reference.

RQ 2: Does content word overlap affect discourse-level text processing by L2 readers?

Research question two refers to the role that the content word overlap plays in L2 reading and is addressed by experiment two in multiple ways. Results are presented here according to early processing measurements (gaze duration) and late processing measurements (regression path duration) for native and nonnative readers and include discussions of the influence of both the quantity and quality of overlap.

In early processing measures, overlap affected nonnative speaker reading but not native speaker reading. For both L2 groups, significantly longer gaze duration times on the target were found when no overlap was present in the passage as compared to when there were four instances of overlap throughout the passage. Thus, the presence of additional overlap facilitated the early processing of the target word. For low L2 learners, differences are also seen between close and distant overlap. Passages with moderate
overlap in adjacent sentences were read faster than passages with moderate overlap separated by a gap of two sentences, albeit this difference in reading time was insignificant. Thus, both quality and quantity of overlap affects second language reading. These effects were not seen for native speakers where all target words were read similarly regardless of overlap quantity or location.

These findings are important to understanding L2 reading as the main effect for overlap was found in passages regardless of passage difficulty. Thus, whether passages were easy or advanced, the absence of overlap impacted the reading of the target word. As a reminder, all target words were of the highest frequency range and therefore should have been easily processed by all readers in the experiment. The main effect of overlap found demonstrates that the early processing of the target word was sensitive to the amount of overlap in the surrounding context. The presence of overlap therefore facilitates processing, regardless of how difficult the surrounding text is, and speeds the early processing of the target word. The low L2 results highlight that when overlap is located in adjacent sentences, the facilitation is increased as opposed to when the overlap is separated by a gap of sentences.

This finding supports the early conclusion of Coady (1979) that low-level readers may rely on contextual information to scaffold their limited lexical and syntactic processing. Since reading requires multiple steps, all demanding resources from a central processor, if certain steps of the reading process can be accomplished with less resources, this frees additional processing capacity for other aspects of reading. The presence of content word overlap facilitates the processing of text such that the reader is able to use more resources for word recognition. Another possible explanation is that the presence of
overlap also increases the activation for the target word; if the reader has seen this word multiple times in the text already, it is more easily accessed (Plummer, Perea, & Rayner, 2014).

Further facilitation of overlap is seen for L2 readers in the late processing measures, here measured as regression path duration. Regression path duration is a measure of reading time once the target has been fixated on for the first time and any regressive reading until a word to the right of the target is fixated. In all passages, target words located in the no overlap condition triggered longer regressive reading than target words in the all or moderate overlap conditions. The longer reading times in the no overlap condition were significant. Longer regression path durations are indicative of longer processing time for integrating the target word into the context of meaning for the reader. Thus, when all or moderate overlap is present, processing is facilitated as evidenced by shorter reading times.

Taken together, experiment two shows that the amount and location of content word overlap impacts L2 reading both by early processing stages and continues to impact nonnative readers as they integrate the word into the context. However, this benefit is not extended to L1 readers. Overlap had no effect on target reading time for L1 readers in this study. These results diverge from the L1 literature on cohesion effects in that there are no overlap effects found for native readers (Britton and Gulgoz, 1991; McNamara et al., 2010). The results of this study do not necessarily mean that readers are not sensitive to cohesion when reading in their native language but that the passages used in this experiment may be too easy for any overlap effects to be required. Since these passages
needed to accommodate low L2 readers, even the most advanced passages may not have caused processing difficulties for L1 readers.

**Research Question three Discussion**

Research question three addresses the role of content overlap in terms of L2 proficiency and whether overlap effects differ as proficiency in L2 increases. Research question three asks:

RQ 3: Do the effects of content word overlap vary based on proficiency of the reader?

The L2 literature has shown diverging reading behaviors as learners advance in their proficiency where advanced readers do not rely on the same strategies as more beginning learners. The results of this study are consistent with this pattern as overlap affected the high L2 group differently than the low L2 group. Initially, global reading measures were examined to get a bigger picture of how the two nonnative speaker proficiency groups exhibited different reading behaviors. Results showed that low L2 readers made more fixations per passage, read more slowly, and comprehended less than the high L2 group, which unsurprisingly aligns with the general consensus of L2 reading literature (Koda, 2005). Thus, the two groups in this experiment show proficiency effects on reading behavior.

In gaze duration measurements, early L2 learners saw consistent incremental increases of about 13ms in reading time as overlap decreased in both quantity and quality. Target words in all overlap passages were read the fastest followed in successive order by close overlap, distant overlap, and then no overlap. Based on these results, low
L2 learners' early processing was affected by both quantity and quality of overlap. However, the more proficient L2 learners only showed significantly longer reading times in the no overlap condition; means for all, close, and distant overlap conditions were approximately the same. Therefore, where low L2 learners were benefitted most by four instances of overlap, two instances of overlap is just as effective when proficiency is more advanced, yet high L2 learners are also still affected by reading passages with no overlap. The only comparable studies in the reading literature to associate with these findings are the L1 differences in global and local cohesion (McNamara et al., 1996).

Proficiency differences also emerge when looking at late processing measures. Both low and high L2 learners have longer regression path duration times in the no overlap condition; however, these differences were only significant for the low L2 group. While high L2 learners have a statistically significant processing cost for accessing the lexical meaning when no overlap is present in the passage, these processing costs do not remain when it comes to integrating the lexical item into the context. These results seem to support the notion that as proficiency develops in the L2, the reliance on content word overlap for facilitating text integration is lessened. This supports the literature that finds more proficient L2 readers are more capable of inference resolution than low L2 readers (Carrell, 1991; Lee and Schallert 1997).

Native speakers performed differently than both nonnative speaker groups. In gaze duration, native speakers saw no significant differences in any condition. Thus, lexical access is swift and unaided by amount of overlap in the passages of experiment two. There also were no significant effects found in regression path duration either. Examining the raw data of both gaze duration and regression path duration find that the
no overlap condition is the fastest condition for native speakers, suggesting that L1 readers benefit from the lack of overlap. Likely the L1 readers have such strong automaticity in lexical access that their working memory is not taxed by a lack of cohesion throughout the text. However, as previously stated, these passages were not written with an L1 reading ability in mind and may be too simplistic and short to capture the effects of content word overlap in L1 reading.

In sum, it appears that low level L2 readers benefit both in early access and sentence integration from increasing amounts of overlap, regardless of passage difficulty. As second language learners advance to higher levels of proficiency, the total absence of overlap still impacts lexical access and to a lesser degree sentence integration, but the differences between four and two instances of overlap no longer affect higher level L2 reading. Despite their high levels of proficiency, high L2 readers still do not read like native speakers who, despite no significant differences, have faster processing times in no overlap conditions than in two or four instances of overlap. As such, it is clear that quantity of overlap influences L2 reading comprehension differently at various proficiency levels.

**Research Question Four Discussion**

The discussion of results above involves influences of cohesion regardless of lexical difficulty, represented in experiment two as passage difficulty levels. The general reading results indicate significant passage difficulty differences, especially for the easy passages. Thus, if both lexical difficulty and overlap difficulty facilitate reading,
research question four is necessary to understand how cohesion and lexical difficulty interact. Research question four states:

RQ 4: Do the effects of content word overlap vary based on difficulty of lexical items in the text?

There was no main interaction for passage difficulty and content word overlap. However, clear, yet insignificant, trends are emerging in the data for low L2 learners in regards to how cohesion impacts reading at different levels of lexical difficulty. In the easy passages in all overlap condition target reading times are approximately the same duration (M=337ms) as those in close overlap passages (327ms) while distant overlap and no overlap conditions yielded longer reading times (362ms and 357ms respectively). Thus, while the presence of four instances of overlap is most beneficial to readers, when passages are of easy lexical difficulty, moderate overlap is sufficient for facilitated processing when it occurs in adjacent sentences. Moderate overlap when separated by a gap of sentences does not result in facilitated target reading times. Also, even in lexically simplistic passages, the absence of overlap results in longer processing in both early and late measures.

This benefit of moderate adjacent overlap is lost when the lexical difficulty of the passages increases as seen in the intermediate and advanced passage results. In both of the proficiency levels where vocabulary was more complex, the all overlap passages result in approximately 40ms shorter reading times for the target than in the close and distant conditions, which in turn were approximately 30ms shorter than the no overlap condition. Again, these gaps were statistically insignificant yet indicative of an emerging
trend. Thus, in more lexically complex passages, low level L2 learners demonstrate that target words in all overlap conditions are read faster than those in moderate overlap passages, and those in no overlap passages result in the slowest target reading times.

Interestingly, high L2 readers show slightly different gaze duration patterns when lexical difficulty is combined with overlap. In lexically complex passages, the general overlap trend for higher proficiency learners was the same: all, close, and distant overlap showed little variation while the no overlap condition target words were read significantly faster. In easy passages, however, target words in the close overlap condition were read faster than those in all conditions, which is comparable to the L1 pattern of overlap. These findings suggest that when passages are lexically less complex, high L2 readers are similar to L1 readers in that adjacent overlap is more beneficial than all overlap. However, high L2 readers differ from L1 readers in that the distant and no overlap conditions still pose a processing cost.

Overall, experiment two results show that for nonnative speakers the absence or presence of overlap has significant effects on reading, both in early processing of the lexical item and in later processing of integrating the word into the discourse. These results are explained by the reading models discussed in Chapter 2 that argue processing resources are limited while reading. If readers are burdened with lexical and syntactic processing, there are less resources available to build inferences and make connections across the text. When referential connections are more explicit, even in short passages, faster early and late processing results due to freed resources.
These results are also significant in light of two lexical difficulty factors. First, all target words, even those in the advanced passages, are of the highest frequency range, meaning they should be easily accessed with a certain degree of automaticity, particularly for the high L2 group. However, the varied reading times in the all and no overlap conditions indicate that lexical access of even high-frequency words is benefitted by the presence of maximal overlap and hindered by the absence of overlap. Secondly, these main effects were found across all passages difficulty levels, but closer analysis revealed that if the lexical difficulty throughout the entire passage is easier, less cohesion is needed for low L2 learners.

Experiment two also helped reveal differences in general reading behavior based on non-native proficiency and compared to native speakers. In general, advanced non-native speakers comprehend about as much of passages ($M = 92\%$) as native speakers ($M = 95\%$); however, advanced nonnative readers take longer and make more fixations per passage to accomplish the same comprehension. Also, the advanced L2 group was more sensitive to the presence of overlap where L1 readers were not. Therefore, advanced L2 readers are engaging in different and more arduous reading behavior to reach the reading level of native readers, and they are benefitted by cohesion to help them do so.

The lack of significant difference in comprehension score but presence of difference in reading behavior is also indicative of the importance of measuring reading using both on-line and off-line procedures. From the off-line measures available in experiment two, we see few differences in comprehension and total reading time across all overlap conditions within participant groups, yet the reading times of the target words and the regressive behavior show that readers are exhibiting different reading behaviors.
to get to the same result. Thus, this highlights the importance of on-line processing investigation in reading research, as different conclusions may have been reached through off-line analyses alone.

5.3 General Discussion

Reading is an effortful task and comprehension is affected by a number of cognitive factors. This dissertation sought to explore the assessment of reading through two different approaches, readability predictions of difficulty and eye-movement analysis, in order to ascertain the influence of cohesion in L2 reading. Results from both experiments support the conclusions from the literature that grammatical and lexical complexity do impact reading comprehension, but content word overlap also plays a significant role in L2 processing of text at multiple stages of development.

A reader's first task when encountering a text is decoding words and retrieving lexical meaning. At this early stage of processing, cohesion shows effects for even high frequency words as seen in the gaze duration results of experiment two. As learners are in the early stages of L2 development, maximal overlap results in the greatest benefits while moderate overlap is slightly less beneficial. More proficient L2 readers do not require as much overlap as the low L2 learners, but they still see consequences in word reading time when no overlap is present. As texts increase in overall lexical complexity, the impacts of cohesion remain significant. For low level L2 learners, the increased lexical difficulty leads to a need for more overlap. Thus, lexical complexity does not impact text understanding in isolation; its effects shift and change as content word overlap is manipulated.
After retrieving the meaning for lexical items, readers then integrate the word meaning into the discourse model of their representation of the text. Experiment two also showed that content word overlap impacts this process for nonnative readers. Low L2 readers incur less rereading time from the target as a sign of faster integration. High L2 readers do not have significant results but raw means show the trend that no overlap impedes this processing through longer rereading times.

In terms of overall reading, experiment two did not yield overall benefits in reading time, total fixations, or comprehension according to overlap; however, the passages used in experiment two were short and contrived of minimal content word overlap aside from the one instance manipulated across passage conditions. When longer and more authentic passages are used, such as those in experiment one, results show how cohesion, when factored in with lexical and syntactic complexity, can be a better predictor of reading difficulty.

Taken together, results from experiments one and two show how cohesion affects lexical access, text integration and overall reading comprehension in line with reading involving multi-level cognitive demands (Koda, 2005). Cohesion helps readers to integrate information into the extended discourse and facilitate reading comprehension, and that L2 reading comprehension is composed of more than just lexical and grammatical competency.
5.4 Limitations and Future Directions

In this section, the limitations of experiments one and two are discussed. Additionally, suggestions are made for future research studies that are not as constrained as the scope of this project.

Experiment One Limitations

There are several factors by which the results of experiment one are limited. The first, and most significant, of these limitations is the limited number of tests for readability analysis. As this study used an existing assessment designed for educational rather than research purposes, the design does not suit itself for statistical analysis. The large number of participants' comprehension scores allow for statistical power in demonstrating that there are differences between tests within a level; however, the connections observed between readability scores and these comprehension results are speculative until confirmed inferentially. Three tests at each level is too small of a number to make an inferential correlation about the pattern between readability predictions and comprehension scores. Future studies should include additional tests within the same level in order to make stronger claims rather than descriptive observations.

Additionally, background knowledge required to read the texts may have affected how learners comprehend the text. Although text topics were chosen for their generality and lack of bias towards particular majors, prior knowledge may have impacted comprehension as seen in other studies (McNamara, Floyd, Best, & Louwerse, 2004). Additionally, cultural knowledge has been shown to be a factor in connecting text
information to prior knowledge (Carrell, 1987). A prior-knowledge task based on the topics of these texts would be informative as to whether cultural gaps or world knowledge was the source of comprehension differences.

Summaries and free-recalls are commonly used as reliable ways of assessing reading comprehension and have been found to result in different comprehension scores than objective questions (Britton & Gülgöz, 1991). Thus, the subjective summary data from this assessment would provide other insights into the role of content word overlap. However, it is important that such subjectively-scored data are reliable through multiple raters and clear rating criteria, which was not possible for the current study. The data from this assessment could be rated a second time to ensure reliability of score and compared to readability predictions.

Lastly, experiment one had a diverse participant pool. To improve research design, the same participants should read all of the tests in the level in order to solidify claims about relative test difficulty. Also, experiment one participants were of many L1 backgrounds; exploring test scores to determine if differences were driven by language background would further inform the L2 reading literature. For instance, certain languages may be more susceptible to content word overlap influences than others.

**Experiment Two Limitations**

Experiment two also had limitations in its large design. There were 12 experimental conditions (four overlap conditions by three difficulty levels); thus, each participant only encountered four of each condition type. Participants with missing data or who did not fixate on the target during the first pass were excluded, which resulted in
loss of data. As such, some numbers were less stable and it was more difficult to identify outlying data, which led to more variability in the means. Future replications should either increase the number of participants or have only three conditions. Doing so would also simplify the analysis and increase statistical power.

Another consequence of the large design was the task was quite taxing for participants. Due to the number of variables being examined, the task needed to be expanded to include a suitable number of trials for each variable, resulting in 48 passages to be read. This resulted in most L2 participants requiring at least one hour to complete the task and likely induced fatigue-related problems. This was compensated for in the current design through the randomization of trials; however, it also meant that the lowest level of L2 participants, beginners, were excluded from the study as they could not complete the task. The results of this study indicate that overlap effects depend on L2 proficiency; thus, a more a study focused solely on low level learners using easier passages should be conducted to see how the lowest level of learners read with and without content word overlap.

Another limitation of experiment two is the variability of the data. Within each L2 group was a significant amount of proficiency variation, which likely led to high variability in the reading data. Restricting the range of vocabulary size scores to produce more homogenous L2 proficiency groups may also produce interesting results. For instance, there were several participants with vocabulary size scores above 15,000; a study focused on comparing these highly proficient learners to L1 would be fruitful for further L1 / L2 reading comparisons.
The passages in experiment two were carefully crafted to manipulate only lexical difficulty with content word overlap in the experimental conditions; syntactic simplicity was left constant across passage difficulty levels and was therefore excluded from any analysis. While this was necessary in the context of experiment two to isolate effects of content word overlap and lexical difficulty, it ignores the role that syntactic complexity plays in text difficulty. Future studies may wish to isolate content word overlap and syntactic complexity to fill in the understanding of how these three measures impact reading.

The passages in experiment two were also written to be short and concise to limit confounding factors. This procedure helped to see overlap effects more clearly but is not reflective of authentic reading texts. That cohesion effects were seen in such short passages is encouraging, as longer passages are more likely to incur cohesion effects as readers have more complex discourse models to build. Also, other cohesion factors are worth exploring such as presence of connective words.

Eye-tracking studies provide numerous measurements, far more than can be described in the scope of this study. For L1 readers, research has shown that the region immediately following the target, term the spillover region, has been a place of processing wrap-up (Rayner et al., 1989). Thus, analysis of the spillover regions in this study, or in studies involving more appropriately challenging texts, may yield interesting results in regards to overlap for L1 readers.

Additionally, number of regressions is often analyzed in eye-movement as a late reading measure to inform text integration difficulties (Clifton, Staub, & Rayner, 2007).
However, this measure may be complicated in L2 reading which is less fluent than L1 reading and prone to more overall rereading. A study designed to specifically examine regressive reading behavior in L2 readers would be influential in understanding other sources of text difficulty.

Finally, background knowledge was only briefly assessed in this study for the purposes of determining that all passages were of comparable familiarity to a general L1 and L2 audience. The L1 literature has shown a strong effect of cohesion for low-knowledge readers and would be of worthwhile investigation for in L2 contexts given the findings of the current study on cohesion in L2 reading (McNamara et al., 2010). A study using more learning-oriented passages and a specific prior-knowledge assessment would enhance the understanding of how these two attributes interact with proficiency to affect L2 reading.

5.5 Pedagogical Implications

This research has several implications for curriculum developers, language assessments, and researchers. Reading in a second language has been shown to be critical for improving overall language proficiency. Thus, ensuring that teachers provide appropriately-leveled texts to stimulate language development. Currently, teachers primarily focus on the lexical and syntactic components of text, but this research suggests that cohesive factors are important to consider as well.

These results are beneficial for teachers as they support the use of a freely-available web tool to analyze texts they are providing to students. Often, teachers are encouraged to use authentic reading texts in the classroom; however, teachers are unsure
of how to classify the reading level of these texts. By using an appropriate readability tool such as Coh-Metrix, they can be more confident that the material they provide to students is in the appropriate readability range.

Reading teachers often rely on pre-leveled material, such as readings in an EFL textbook; however, this is not a guarantee that all texts within one textbook are of comparable reading level. As seen in experiment one, all of the texts for the assessment came from the same textbook designed and published for specific levels of learner. Yet for some of the levels, comprehension differences emerged despite all three texts coming from the same textbook. Had curriculum developers of that series used appropriate readability tools to ensure continuity of text difficulty, there would have been less variability in performance as with the levels that did not have readability differences.

These results are also significant for assessment writers. Without appropriately controlling text difficulty, it can be impossible to assess reading comprehension accurately. For instance, the EFL language program from which the scores for experiment one were obtained has used the assessments for multiple years without recognizing the gap in comprehension performance on some of the tests. Thus, it is possible that certain readers were held back for a lower comprehension score without the language program realizing the readability of the text was different from the other tests. For high-stakes proficiency assessments, like TOEFL or IELTS, it is even more critical to ensure stability of passage difficulty.

Lastly, researchers investigating L2 reading must consider cohesion when designing texts for experimental purposes. Experiment two indicated how in closely
controlled passages cohesion can result in processing differences. Conclusions made in experiments that did not control for these measures could be flawed and cohesion may be a confounding factor that is overlooked.

5.6 Conclusions

The goal of this dissertation is to shed light on the role of cohesion in L2 reading through two different means of assessing comprehension. Specifically, experiment one demonstrated how a cognitively-based readability formula that involves cohesion in addition to lexical and syntactic complexity is more predictive of comprehension performance difference than traditional readability formulas that only consider surface-level features of a text. Experiment two supported these findings by investigating cohesion and its individual contributions in L2 reading. Experiment two shows that both amount and quality of overlap affect low level L2 readers in their processing of text. The benefits for high L2 readers shift as their lexical competency increases but benefits of cohesion are still evident. Experiment two also indicates how off-line and on-line processing measures are both needed to identify reading behavior and further explore the differences between native and non-native readers.

This research is the first to explore content word overlap in L2 reading using on-line processing measures and is just an initial exploration into its effects. Additional research is needed to further advance our understanding of the complex processes involved in reading.
REFERENCES


1. Read the article and write a summary.
2. Turn in your summary and get questions 2-8 from your teacher.
3. Answer questions 2-8. Submit article and questions to your teacher. Make sure your name is on each page.

**The Aral Sea: An Environmental Crisis**

(1) For decades, environmental scientists have warned of the ecological damage that results from poor planning to stimulate and develop the economy. Reasonable solutions for more employment, housing, and food come at the expense of clearing the land, and don’t consider the future consequences. One of the world’s most extreme cases of environmental disasters is the Aral Sea region, located between Kazakhstan and Uzbekistan. This region is a clear example of the damage that poorly planned human economic activity can have on the environment.

(2) In the 1950s, this sea was the world’s fourth largest lake. Two large rivers, the Amu Darya and the Syr Darya, flowed into the sea. Although the land around the sea and the rivers was dry and desert-like, the region still supported a wide diversity of wildlife, including 500 species of birds, 200 species of mammals, and numerous species of fish. In the late 1950s, the Soviet Union government decided to develop agriculture to improve the economy in this dry area by using water from the two rivers. The government built huge farms, digging long paths to bring water from the rivers. By 1965, the Amu Darya River was providing water across 7 million acres (28,000 square kilometers) of wheat and cotton. These crops flourished with the help of huge amounts of chemical pesticides and fertilizers.

(3) In the short term, this plan was successful. In the early 1960s, the farms and the fishing industries saw great economic benefits. Each year, more than 40,000 tons of fish were caught and transported by rail to Moscow. Thousands of people were employed in the boats, processing plants, and railroad yards. The irrigation system made the once dry desert fertile, encouraging more farms to be established.
(4) Within a few years, however, it was clear that this agricultural effort was not only unsustainable, but was, in fact, causing devastating damage to the environment. The Amu Darya River was reduced drastically, to about half its original size by the early 1990s. Eventually, as it receded, this river was cut off from the Aral Sea. Moreover, the reduction in volume and water flow led to increased salt levels. Salt content climbed from 10 to 30 grams per liter, almost as high as the oceans. Native species of fish died out, and the commercial fishing industry collapsed.

(5) Because of the enormous size and complexity of the crisis, practical responses were slow to appear. After the fall of the Soviet Union, however, the World Bank stepped in, and together with the newly independent Kazakhstan, began to tackle the problem by funding extensive projects aimed at saving the northern part of the Aral Sea. By 2005, they had built a dam intended to raise the water level of the North Aral by about 13 feet.

(6) This would return the water to its previous salt level, and therefore, allow native fish to repopulate. The results were promising. The North Aral Sea grew by 20 percent. The salt levels fell to almost the 1960 level, and freshwater fish returned. The fishing industry slowly rebounded, with 1,500 tons caught in 2008. Today, fish-processing plants are operating again, providing much needed employment. However, the condition of the South Aral Sea has continued to worsen, and scientists predict that it will be completely dry by 2020.
1) Write a summary for this text. *Do NOT copy the text; put the answer in your own words.*

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
Use the text to answer the following questions.

2) What is the main idea of this article?
   a. The environmental effects are worth the cost of building the economy.
   b. Environmental consequences should be considered when developing an economy.
   c. The North Aral Sea has recovered from the damage caused by poorly planned economic activity.
   d. People should be more concerned about the environment than the economy.

3) In paragraph (6), what does this refer to?
   a. the northern part of the Aral Sea
   b. 13 ft
   c. the water level
   d. salt content

4) Which of the following is TRUE about the Aral Sea before the 1950s?
   a. There was a dam to help raise the water level.
   b. The land didn’t have a lot of animals because it was a desert.
   c. There were a lot of government farms nearby.
   d. None of the above

5) In paragraph (6), what is the best meaning for rebounded?
   a. recovered
   b. grew
   c. slowed down
   d. operated

6) What statement would the author agree with?
   a. Pesticides will help improve the economy of a nation by making crops grow faster.
   b. The use of trains was inefficient for the sale of fish in Moscow.
   c. A solution that fixes a current problem could cause unknown problems in the future.
   d. The Soviet Union’s response was appropriate for the crisis.

7) What is the meaning of drastically in paragraph (4)?

8) How long did it take to respond to the crisis in the Aral Sea? Why? Do NOT copy the text; put the answer in your own words.
APPENDIX B

EXPERIMENT TWO PASSAGE METRICS AND FAMILIARITY RATINGS

Table information includes: passage number, number of characters in target word, frequency of target (taken from the Brigham Young University COCA Corpus), topic, and familiarity survey results for both native and nonnative speakers, where one indicates little knowledge and six indicates great knowledge.

<table>
<thead>
<tr>
<th>#</th>
<th>Char</th>
<th>Target Word</th>
<th>Target Freq.</th>
<th>Topic</th>
<th>Familiarity Survey Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>children</td>
<td>820</td>
<td>Role of children in Inuit communities</td>
<td>NS 1.5 NNS 2.1</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>school</td>
<td>125</td>
<td>Historical changes in educational university policies</td>
<td>NS 2.5 NNS 2.1</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>rule</td>
<td>633</td>
<td>A 10,000 year old tree in Sweden</td>
<td>NS 1.36 NNS 2.13</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>strategy</td>
<td>845</td>
<td>Strategies of discount airlines</td>
<td>NS 2.4 NNS 2.9</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>practice</td>
<td>564</td>
<td>Estonian triplets who are Olympic athletes</td>
<td>NS 1.39 NNS 2.07</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>problem</td>
<td>171</td>
<td>How airlines are trying to fix security line length</td>
<td>NS 2.69 NNS 2.71</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>store</td>
<td>700</td>
<td>Innovative markets in Indonesia</td>
<td>NS 1.3 NNS 1.5</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>solution</td>
<td>1453</td>
<td>Strategies of disgruntled employees in cold offices</td>
<td>NS 2.3 NNS 2.6</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>vacation</td>
<td>2638</td>
<td>Reasons why you should vacation in the Dominican Republic</td>
<td>NS 2.6 NNS 2</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>country</td>
<td>166</td>
<td>Reasons why Canada is a good place to study abroad</td>
<td>NS 2.4 NNS 2.7</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>business</td>
<td>246</td>
<td>The business that makes all the clothes of the soldiers in the world</td>
<td>NS 2 NNS 2.07</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>location</td>
<td>1437</td>
<td>History of Pamukkale</td>
<td>NS 1.3 NNS 2.14</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
<td>report</td>
<td>476</td>
<td>Reactions of businesses and markets to Brexit</td>
<td>NS 2.1 NNS 2.1</td>
</tr>
<tr>
<td>14</td>
<td>6</td>
<td>desire</td>
<td>1556</td>
<td>How American students try to obtain competitive scholarships</td>
<td>NS 4.1 NNS 2.42</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>belief</td>
<td>1382</td>
<td>Cultural differences on the superstition of black cats</td>
<td>NS 2.6 NNS 3.1</td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>category</td>
<td>1465</td>
<td>Research on the difference between starting school at age 4 or at age 5</td>
<td>NS 2.5 NNS 2.8</td>
</tr>
</tbody>
</table>

**Avg** 6.9 916 2.19 2.34
Table information includes: passage number, number of characters in target word, frequency of target (taken from the Brigham Young University COCA Corpus), topic, and familiarity survey results for both native and nonnative speakers, where one indicates little knowledge and six indicates great knowledge.

<table>
<thead>
<tr>
<th>#</th>
<th>Char.</th>
<th>Target Word</th>
<th>Target Freq.¹</th>
<th>Topic</th>
<th>Familiarity Survey Results²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>connection</td>
<td>1425</td>
<td>Common activities that cause cancer</td>
<td>3.82</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>mistake</td>
<td>1434</td>
<td>Problems of planning a wedding</td>
<td>3.21</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>beginning</td>
<td>1202</td>
<td>The origin of common businesses</td>
<td>3.23</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>possibility</td>
<td>1197</td>
<td>Odds of winning the lottery</td>
<td>3.62</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>score</td>
<td>1071</td>
<td>Subjective scoring in some Olympic sports</td>
<td>2.78</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>evidence</td>
<td>606</td>
<td>Beyonce's advertising contract with Pepsi</td>
<td>2.29</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>period</td>
<td></td>
<td>History of immigration in the US</td>
<td>3.70</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>career</td>
<td>741</td>
<td>Salaries of baristas.</td>
<td>2.14</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>figure</td>
<td>552</td>
<td>Climate change chart by Michael Mann</td>
<td>1.57</td>
</tr>
<tr>
<td>0</td>
<td>10</td>
<td>difference</td>
<td>496</td>
<td>History of diversity in the workplace</td>
<td>3.45</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>status</td>
<td>1217</td>
<td>Ranks in the US army</td>
<td>3.20</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>change</td>
<td>256</td>
<td>Influence of Peter the Great on schools</td>
<td>1.79</td>
</tr>
<tr>
<td>13</td>
<td>8</td>
<td>facility</td>
<td>1301</td>
<td>History of the building of the Coca Cola Company</td>
<td>2.40</td>
</tr>
<tr>
<td>14</td>
<td>10</td>
<td>experience</td>
<td>420</td>
<td>First hiker of the Appalachian Trial</td>
<td>1.97</td>
</tr>
<tr>
<td>15</td>
<td>9</td>
<td>knowledge</td>
<td>840</td>
<td>What ombudsmen do in the workplace</td>
<td>1.89</td>
</tr>
<tr>
<td>16</td>
<td>7</td>
<td>variety</td>
<td>1172</td>
<td>History of the store Macy's</td>
<td>1.78</td>
</tr>
<tr>
<td>Avg</td>
<td>7.3</td>
<td></td>
<td>947</td>
<td></td>
<td>2.68</td>
</tr>
</tbody>
</table>
Table information includes: passage number, number of characters in target word, frequency of target (taken from the Brigham Young University COCA Corpus), topic, and familiarity survey results for both native and nonnative speakers, where one indicates little knowledge and six indicates great knowledge.

<table>
<thead>
<tr>
<th>#</th>
<th>Char.</th>
<th>Target Word</th>
<th>Target Freq.</th>
<th>Topic</th>
<th>NS</th>
<th>NNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>require</td>
<td>1895</td>
<td>Teen Driving Laws</td>
<td>4.31</td>
<td>3.20</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>acceptance</td>
<td>1232</td>
<td>Minor marriage laws</td>
<td>2.18</td>
<td>2.41</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>difficulty</td>
<td>1693</td>
<td>Difficulties in starting a small business</td>
<td>3.40</td>
<td>2.67</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>fear</td>
<td>986</td>
<td>Fears of first-time parents</td>
<td>3.61</td>
<td>2.75</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>enjoyment</td>
<td>321</td>
<td>Choosing good employees</td>
<td>3.71</td>
<td>2.71</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>fight</td>
<td>1569</td>
<td>Roommate fights</td>
<td>3.68</td>
<td>3.01</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>mission</td>
<td>1096</td>
<td>Boldness of business majors</td>
<td>2.70</td>
<td>2.18</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>idea</td>
<td>311</td>
<td>Creation of Amazon.com</td>
<td>2.43</td>
<td>2.51</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>device</td>
<td>1512</td>
<td>New inventions to prevent ocean pollution</td>
<td>2.88</td>
<td>2.40</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>money</td>
<td>582</td>
<td>Use of Presidential campaign money</td>
<td>3.39</td>
<td>2.01</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>victory</td>
<td>1467</td>
<td>The creation of a roller coaster in Ohio</td>
<td>1.65</td>
<td>1.90</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
<td>purpose</td>
<td>1015</td>
<td>Story of Bethany Hamilton</td>
<td>2.60</td>
<td>2.00</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>crowd</td>
<td>1274</td>
<td>What happened during a violent Trump rally</td>
<td>3.51</td>
<td>2.14</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>leaving</td>
<td>150</td>
<td>The departure of Richard Nixon</td>
<td>2.99</td>
<td>1.66</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>search</td>
<td>1505</td>
<td>The search to find the Holy Grail</td>
<td>2.29</td>
<td>2.04</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>supply</td>
<td>1499</td>
<td>The resources of the Cambridge Library</td>
<td>2.50</td>
<td>2.37</td>
</tr>
<tr>
<td>Avg</td>
<td>6.8</td>
<td></td>
<td>1009</td>
<td></td>
<td>2.98</td>
<td>2.37</td>
</tr>
</tbody>
</table>
APPENDIX C

EXPERIMENT TWO PASSAGES IN OVERLAP CONDITIONS


The Inuit people in Alaska protect their culture in many ways. They take care of their babies as a connected community. As the kids get older, everyone in the village works together with other families. For example, all of the fathers take their sons to learn how to fish together. The community sees children as their culture’s future.

WC: 58


The Inuit people in Alaska protect their culture in many ways. They take care of their babies as a connected community. As the kids get older, everyone in the village works together with other families. For example, all of the fathers take their children to learn how to fish together. The community sees children as their culture’s future.


The Inuit people in Alaska protect their culture in many ways. They take care of their children as a connected community. As the kids get older, everyone in the village works together with other families. For example, all of the fathers take their sons to learn how to fish together. The community sees children as their culture’s future.
The Inuit people in Alaska protect their culture in many ways. They take care of their children as a connected community. As the children get older, everyone in the village works together with other families. For example, all of the fathers take their children to learn how to fish together. The community sees children as their culture’s future.

Educational practices have changed a lot over the last 20 years. Many colleges give more A and B grades than ever before. People say universities are doing this to keep students happy, like a business. Instead of giving students the scores they earned, the institutions give the grades students want. If this pattern continues, the graduates from the schools will be less prepared.
the grades students want. If this pattern continues, the graduates from the schools will be less prepared.

Beg. Passage 2 (D): school - university - institution - school

Educational practices have changed a lot over the last 20 years. Many schools give more A and B grades than ever before. People say universities are doing this to keep students happy, like a business. Instead of giving students the scores they earned, the colleges give the grades students want. If this pattern continues, the graduates from the schools will be less prepared.

Beg. Passage 2 (A): school - school - school - school

Educational practices have changed a lot over the last 20 years. Many schools give more A and B grades than ever before. People say schools are doing this to keep students happy, like a business.

Instead of giving students the scores they earned, the schools give the grades students want. If this pattern continues, the graduates from the schools will be less prepared.

Beg. Passage 2 (No Target): school - school - school - school

Educational practices have changed a lot over the last 20 years. Many give more A and B grades than ever before. People say are doing this to keep students happy, like a business.

Instead of giving students the scores they earned, the give the grades students want. If this pattern continues, the graduates from the will be less prepared.
Researchers have found a 9,550 year old tree in Sweden. The country has made a regulation in order to protect the tree. The rule makes it illegal to cut down the tree. The law also limits the number of tourists who can see the tree at one time. The tree will last a lot longer thanks to the order from the government.
Researchers have found a 9,550 year old tree in Sweden. The country has made an order to protect the tree. The makes it illegal to cut down the tree. The also limits the number of tourists who can see the tree at one time. The tree will last a lot longer thanks to the from the government.

Discount airlines, like Spirit or Ryan Air, are finding new ways to make money. The companies' method is to charge people extra for other benefits. Their approach includes offering the lowest prices on flights. After they get the customers, their plan is to add extra fees for baggage. The airlines have made a lot of money with their strategy in the last year.
offering the lowest prices on flights. After they get the customers, their plan is to add extra fees for baggage. The airlines have made a lot of money with their strategy in the last year.

Beg. Passage 4(A): strategy -- strategy -- strategy -- strategy

Discount airlines, like Spirit or Ryan Air, are finding new ways to make money. The companies' strategy is to charge people extra for other benefits. Their strategy includes offering the lowest prices on flights. After they get the customers, their strategy is to add extra fees for baggage. The airlines have made a lot of money with their strategy in the last year.

Beg. Passage 4(No Target): strategy -- strategy -- strategy -- strategy

Discount airlines, like Spirit or Ryan Air, are finding new ways to make money. The companies' is to charge people extra for other benefits. Their includes offering the lowest prices on flights. After they get the customers, their is to add extra fees for baggage. The airlines have made a lot of money with their in the last year.

Beg. Passage 5(N): training -- exercise -- preparation -- practice

Three identical sisters, triplets, will compete in the Olympics for the first time. The sisters, who are from Estonia, are finishing their training together. Their parents encouraged them to do their Olympic exercises together. Although they do their preparation as a team, they will compete to win first place. They like both the practice and competition as a family.

WC: 59
Three identical sisters, triplets, will compete in the Olympics for the first time. The sisters, who are from Estonia, are finishing their training together. Their parents encouraged them to do their Olympic exercises together. Although they do their practice as a team, they will compete to win first place. They like both the practice and competition as a family.
to do their Olympic together. Although they do their as a team, they will compete to win first place. They like both the and competition as a family.

Beg. Passage 6(N): challenge -- issue -- concern -- **problem**

It takes a long time to go through security at an airport these days. Airports want to fix the **challenge** because people keep missing their flights. Passengers are mad because of the **issue** and want changes. Security thinks they can fix the **concern** with new machines. They will use the new machines in 2016 and see if the **problem** gets fixed this year.

WC:63

Beg. Passage 6(C): challenge -- issue -- **problem** -- **problem**

It takes a long time to go through security at an airport these days. Airports want to fix the **challenge** because people keep missing their flights. Passengers are mad because of the **issue** and want changes. Security thinks they can fix the **problem** with new machines. They will use the new machines in 2016 and see if the **problem** gets fixed this year.

Beg. Passage 6(D): **problem** -- issue -- challenge -- **problem**

It takes a long time to go through security at an airport these days. Airports want to fix the **problem** because people keep missing their flights. Passengers are mad because of the **issue** and want changes. Security thinks they can fix the **challenge** with new machines. They will use the new machines in 2016 and see if the **problem** gets fixed this year.
Beg. Passage 6(A): problem -- problem -- problem -- problem

It takes a long time to go through security at an airport these days. Airports want to fix the problem because people keep missing their flights. Passengers are mad because of the problem and want changes. Security thinks they can fix the problem with new machines. They will use the new machines in 2016 and see if the problem gets fixed this year.

Beg. Passage 6(No Target): problem -- problem -- problem -- problem

It takes a long time to go through security at an airport these days. Airports want to fix the problem because people keep missing their flights. Passengers are mad because of the problem and want changes. Security thinks they can fix the problem with new machines. They will use the new machines in 2016 and see if the problem gets fixed this year.

Beg. Passage 7(N): market -- shop -- stand -- store

In the country of Indonesia, you can buy your groceries on the water. You travel by boat through the water to the shop and look for the food you want. The name of the market is Muara Kuin, which means "floating." They have all of the normal fruits and vegetables at the stand. They make a lot of money at this store because of the tourists.

WC: 66

Beg. Passage 7(C): market -- shop -- store -- store

In the country of Indonesia, you can buy your groceries on the water. You travel by boat through the water to the shop and look for the food you want. The name of the market is
Muara Kuin, which means "floating." They have all of the normal fruits and vegetables at the store. They make a lot of money at this store because of the tourists.

Beg. Passage 7(D): store -- shop -- market -- store

In the country of Indonesia, you can buy your groceries on the water. You travel by boat through the water to the shop and look for the food you want. The name of the store is Muara Kuin, which means "floating." They have all of the normal fruits and vegetables at the market. They make a lot of money at this store because of the tourists.

Beg. Passage 7(A): store -- store -- store -- store

In the country of Indonesia, you can buy your groceries on the water. You travel by boat through the water to the store and look for the food you want. The name of the store is Muara Kuin, which means "floating." They have all of the normal fruits and vegetables at the store. They make a lot of money at this store because of the tourists.

Beg. Passage 7(No Target): store -- store -- store -- store

In the country of Indonesia, you can buy your groceries on the water. You travel by boat through the water to the store and look for the food you want. The name of the store is Muara Kuin, which means "floating." They have all of the normal fruits and vegetables at the store. They make a lot of money at this store because of the tourists.

Beg. Passage 8 (N): plan -- method -- approach -- strategy

In 2015, 56% of employees thought their office was too cold. These cold employees usually develop a plan to make their office warmer. Their method is to make a lot of angry signs for the boss to see. The employees also sent e-mails to the office as a part of
their approach. Unfortunately, they were unsuccessful and are looking for a new strategy to warm the office.

WC:67

Beg. Passage 8 (C):  plan -- method -- strategy -- strategy

In 2015, 56% of employees thought their office was too cold. These cold employees usually develop a plan to make their office warmer. Their method is to make a lot of angry signs for the boss to see. The employees also sent e-mails to the office as a part of their strategy. Unfortunately, they were unsuccessful and are looking for a new strategy to warm the office.

Beg. Passage 8 (D):  strategy -- method -- plan -- strategy

In 2015, 56% of employees thought their office was too cold. These cold employees usually develop a strategy to make their office warmer. Their method is to make a lot of angry signs for the boss to see. The employees also sent e-mails to the office as a part of their plan. Unfortunately, they were unsuccessful and are looking for a new strategy to warm the office.

Beg. Passage 8 (A):  strategy -- strategy -- strategy -- strategy

In 2015, 56% of employees thought their office was too cold. These cold employees usually develop a strategy to make their office warmer. Their strategy is to make a lot of angry signs for the boss to see. The employees also sent e-mails to the office as a part of their strategy. Unfortunately, they were unsuccessful and are looking for a new strategy to warm the office.
Beg. Passage 8 (No Target): **strategy** -- **strategy** -- **strategy** -- **strategy**

In 2015, 56% of employees thought their office was too cold. These cold employees usually develop a to make their office warmer. Their is to make a lot of angry signs for the boss to see. The employees also sent e-mails to the office as a part of their. Unfortunately, they were unsuccessful and are looking for a new to warm the office.

Beg. Passage 9(N): **vacation** -- **holiday** -- **trip** -- **visit**

The next country you should travel to is the Dominican Republic. There are beautiful beaches and relaxing areas to go to on your **vacation**. The **holiday** will give you beautiful pictures and memories forever. You can try new food and visit different places on your **trip**. You can visit many places, but the best for your **visit** is the Dominican Republic.

Beg. Passage 9(C): **vacation** -- **holiday** -- **visit** -- **visit**

The next country you should travel to is the Dominican Republic. There are beautiful beaches and relaxing areas to go to on your **vacation**. The **holiday** will give you beautiful pictures and memories forever. You can try new food and visit different places on your **visit**. You can visit many places, but the best for your **visit** is the Dominican Republic.

Beg. Passage 9(D): **visit** -- **holiday** -- **vacation** -- **visit**

The next country you should travel to is the Dominican Republic. There are beautiful beaches and relaxing areas to go to on your **visit**. The **holiday** will give you beautiful pictures and memories forever. You can try new food and visit different places on your **vacation**. You can visit many places, but the best for your **visit** is the Dominican Republic.
The next country you should travel to is the Dominican Republic. There are beautiful beaches and relaxing areas to go to on your visit. The visit will give you beautiful pictures and memories forever. You can try new food and visit different places on your visit. You can visit many places, but the best for your visit is the Dominican Republic.

Canadian universities are great places for students to study abroad. There are many top universities in this land for several different majors. The nation has very cheap housing which is good for students. The bigger cities in the state have good entertainment, like festivals and nightclubs. Students looking to get out of their country should consider the country for a semester.
festivals and nightclubs. Students looking to get out of their country should consider the country for a semester.

Beg. Passage 10(D): **country -- land -- nation -- country**

Canadian universities are great places for students to study abroad. There are many top universities in this country for several different majors. The nation has very cheap housing which is good for students. The bigger cities in the land have good entertainment, like festivals and nightclubs. Students looking to get out of their country should consider the country for a semester.

Beg. Passage 10(A): **country -- country -- country -- country**

Canadian universities are great places for students to study abroad. There are many top universities in this country for several different majors. The country has very cheap housing which is good for students. The bigger cities in the country have good entertainment, like festivals and nightclubs. Students looking to get out of their country should consider the country for a semester.

Beg. Passage 10(no targets)

Canadian universities are great places for students to study abroad. There are many top universities in this for several different majors. The has very cheap housing which is good for students. The bigger cities in the have good entertainment, like festivals and nightclubs. Students looking to get out of their country should consider the for a semester.
Do you know who makes the clothes of many soldiers in the world? A company called Wyedean makes uniforms for 60 different countries' armies. The firm started in 1852 and is located in a small town in England. Many people don't even know the organization is real. With more countries looking for uniforms, the business hopes to grow.
Beg. Passage 11 (No targets)

Do you know who makes the clothes of many soldiers in the world? A called Wyedean makes uniforms for 60 different countries' armies. The started in 1852 and is located in a small town in England. Many people don't even know the is real. With more countries looking for uniforms, the hopes to grow.

Beg. Passage 12 (N): site -- place -- spot -- location

Pamukkale, located in southern Turkey, is very popular with tourists. The site was founded in 190 B.C. by the ancient Romans. There are over 50 natural swimming pools at this place. Nowadays, hundreds of tourists visit the spot for the healing abilities of the pools. There is a lot of beauty and culture at the location for people to see.

WC: 60

Beg. Passage 12 (C): site -- place -- location -- location

Pamukkale, located in southern Turkey, is very popular with tourists. The site was founded in 190 B.C. by the ancient Romans. There are over 50 natural swimming pools at this place. Nowadays, hundreds of tourists visit the location for the healing abilities of the pools. There is a lot of beauty and culture at the location for people to see.

Beg. Passage 12 (D): location -- place -- site -- location

Pamukkale, located in southern Turkey, is very popular with tourists. The location was founded in 190 B.C. by the ancient Romans. There are over 50 natural swimming pools at this place. Nowadays, hundreds of tourists visit the site for the healing abilities of the pools. There is a lot of beauty and culture at the location for people to see.
Pamukkale, located in southern Turkey, is very popular with tourists. The location was founded in 190 B.C. by the ancient Romans. There are over 50 natural swimming pools at this location. Nowadays, hundreds of tourists visit the location for the healing abilities of the pools. There is a lot of beauty and culture at the location for people to see.

On June 23, 2016 the British people voted to leave the European Union. People around the world reacted to the news in many ways. European leaders were nervous other countries would want to leave because of the information. The markets were worried by the knowledge, and the British Pound lost a lot of value. In general, people don't know what life will look like now that this report has been released.
the report, and the British Pound lost a lot of value. In general, people don't know what life will look like now that this report has been released.

Beg. Passage 13(C) : report -- news -- information -- report

On June 23, 2016 the British people voted to leave the European Union. People around the world reacted to the report in many ways. European leaders were nervous other countries would want to leave because of the information. The markets were worried by the news, and the British Pound lost a lot of value. In general, people don't know what life will look like now that this report has been released.

Beg. Passage 13(A) : report -- report -- report -- report

On June 23, 2016 the British people voted to leave the European Union. People around the world reacted to the report in many ways. European leaders were nervous other countries would want to leave because of the report. The markets were worried by the report, and the British Pound lost a lot of value. In general, people don't know what life will look like now that this report has been released.

Beg. Passage 13(No Target) : information -- news -- knowledge -- report

On June 23, 2016 the British people voted to leave the European Union. People around the world reacted to the in many ways. European leaders were nervous other countries would want to leave because of the. The markets were worried by the, and the British Pound lost a lot of value. In general, people don't know what life will look like now that this has been released.
Scholarships are very competitive among high school students in the U.S. Because of the high tuition costs, most high school students have a wish to have a scholarship. About seven billion dollars are set aside each year to help students with their dream. Students spend hours online looking for opportunities for their hope. Thanks to their hard work, over one million students got their desire last year.
Beg. Passage 14(A): desire -- desire -- desire -- desire

Scholarships are very competitive among high school students in the U.S. Because of the high tuition costs, most high school students have a desire to have a scholarship. About seven billion dollars are set aside each year to help students with their desire. Students spend hours online looking for opportunities for their desire. Thanks to their hard work, over one million students got their desire last year.

Beg. Passage 14(No Target): wish -- dream -- hope -- desire

Scholarships are very competitive among high school students in the U.S. Because of the high tuition costs, most high school students have a desire to have a scholarship. About seven billion dollars are set aside each year to help students with their desire. Students spend hours online looking for opportunities for their desire. Thanks to their hard work, over one million students got their desire last year.

Beg. Passage 15(N): idea -- belief -- view -- thought

Cultures have a variety of superstitions throughout the world. One common idea in the USA, Europe, and Asia is that black cats are bad luck. Many say the belief started in Egypt over five thousand years ago. People who have this view run away from black cats in the street. Other cultures disagree and say this thought is for children only.

WC:61

Beg. Passage 15(C): idea -- belief -- thought -- thought

Cultures have a variety of superstitions throughout the world. One common idea in the USA, Europe, and Asia is that black cats are bad luck. Many say the belief started in
Egypt over five thousand years ago. People who have this **thought** run away from black cats in the street. Other cultures disagree and say this **thought** is for children only.

**Beg. Passage 15(D) : thought -- belief -- idea -- thought**

Cultures have a variety of superstitions throughout the world. One common **thought** in the USA, Europe, and Asia is that black cats are bad luck. Many say the **belief** started in Egypt over five thousand years ago. People who have this **idea** run away from black cats in the street. Other cultures disagree and say this **thought** is for children only.

**Beg. Passage 15(A) : thought -- thought -- thought -- thought**

Cultures have a variety of superstitions throughout the world. One common **thought** in the USA, Europe, and Asia is that black cats are bad luck. Many say the **thought** started in Egypt over five thousand years ago. People who have this **thought** run away from black cats in the street. Other cultures disagree and say this **thought** is for children only.

**Beg. Passage 15(No Target) : idea -- belief -- view -- thought**

Cultures have a variety of superstitions throughout the world. One common in the USA, Europe, and Asia is that black cats are bad luck. Many say the started in Egypt over five thousand years ago. People who have this **run away from black cats in the street. Other cultures disagree and say this **thought** is for children only.

**Beg. Passage 16(N) : group -- type -- kind -- category**

Educational research says children who go to school at age 4 are better off than those starting at age 5. In 2015, 49% of children were in this **group** in the US. This **type** does better in reading than children who start school at age 5. Researchers also claim this **kind**
will have 23% higher math skills. Based on the research, schools encourage this category to all parents.

WC: 68

Beg. Passage 16(C) : group -- type -- category -- category

Educational research says children who go to school at age 4 are better off than those starting at age 5. In 2015, 49% of children were in this group in the US. This type does better in reading than children who start school at age 5. Researchers also claim this category will have 23% higher math skills. Based on the research, schools encourage this category to all parents.

Beg. Passage 16(D) : category -- type -- group -- category

Educational research says children who go to school at age 4 are better off than those starting at age 5. In 2015, 49% of children were in this category in the US. This type does better in reading than children who start school at age 5. Researchers also claim this group will have 23% higher math skills. Based on the research, schools encourage this category to all parents.

Beg. Passage 16(A) : category -- category -- category -- category

Educational research says children who go to school at age 4 are better off than those starting at age 5. In 2015, 49% of children were in this category in the US. This category does better in reading than children who start school at age 5. Researchers also claim this category will have 23% higher math skills. Based on the research, schools encourage this category to all parents.
Educational research says children who go to school at age 4 are better off than those who start at age 5. In 2015, around 49% of children were in this in the United States. This does better in reading than children who start school at age 5. Researchers also claim that this will have 23% higher math skills. Based on the research, schools encourage this to parents in their area.

**Intermediate Passages + Comprehension Questions**

**Int. Passage 1(N):** link – association—interaction – **connection**

Scientists comprehend more about cancer causes than ever before. Several risk factors can now be identified as **links** to different types of cancer. For instance, smoking and tobacco have long been known to have **associations** to lung cancer. Other **interactions** between the deadly disease and normal habits are being discovered daily. The fight against cancer will be won as these **connections** continue to be discovered.

WC: 65

**True** or False: Scientists continue to find new cancer causes each day.

True or **False**: This passage is about different treatments for cancer.

**Int. Passage 1(C):** link – association—**connection** – **connection**

Scientists comprehend more about cancer causes than ever before. Several risk factors can now be identified as **links** to different types of cancer. For instance, smoking and tobacco have long been known to have **associations** to lung cancer. Other **connections**
between the deadly disease and normal habits are being discovered daily. The fight against cancer will be won as these connections continue to be discovered.

Int. Passage 1(D): connections – association—links – connection

Scientists comprehend more about cancer causes than ever before. Several risk factors can now be identified as connections to different types of cancer. For instance, smoking and tobacco have long been known to have associations to lung cancer. Other links between the deadly disease and normal habits are being discovered daily. The fight against cancer will be won as these connections continue to be discovered.

Int. Passage 1(A): connections – connections—connections – connection

Scientists comprehend more about cancer causes than ever before. Several risk factors can now be identified as connections to different types of cancer. For instance, smoking and tobacco have long been known to have connections to lung cancer. Other connections between the deadly disease and normal habits are being discovered daily. The fight against cancer will be won as these connections continue to be discovered.

Int. Passage 1(No Target): connections – connections—connections – connection

Scientists understand more about cancer than ever before. Several risk factors can now be identified as to cancer. For example, smoking has long been known to have to lung cancer. Other are being discovered daily. The fight against cancer will be won as these continue to be discovered.
People spend more time organizing their wedding than any other life event. However, even the best made plans can involve errors. There are a number of possible disasters like forgetting something crucial at home. Most faults come from attempting to manage too many things at once. When planning a wedding, make sure to tackle tasks one at a time to avoid mistakes that are preventable.
too many things at once. When planning a wedding, make sure to tackle tasks one at a time to avoid mistakes that are preventable.

Int. Passage 2(A): mistakes – mistakes—mistakes – mistake People spend more time organizing their wedding than any other life event. However, even the best made plans can involve mistakes. There are a number of possible mistakes like forgetting something crucial at home. Most mistakes come from attempting to manage too many things at once. When planning a wedding, make sure to tackle tasks one at a time to avoid mistakes that are preventable.

Int. Passage 2(No Target): mistakes – mistakes—mistakes – mistake

People spend more time planning their wedding than they do any other event in their life. However, even the best made plans can end in. There are a number of possible like forgetting something at home or scheduling a company for the wrong day. Most come from trying to do too many things at once. When planning a wedding, make sure to take things one at a time to avoid that could be prevented.

Int. Passage 3(N): foundation – origin—creation – beginning Successful companies weren’t always prosperous immediately. Many popular corporations today have a noteworthy story about their foundation. Some companies had an origin involving conflict when they split from an existing company. Others began with a very modest creation, like Mark Zuckerberg beginning Facebook in his dorm room. We can better understand the present success of a company by looking at its beginning several years ago.

WC: 65
True or False: Facebook was created in a dorm room at a university.

True or False: This passage is about companies that are successful right away.

Int. Passage 3(C): foundation – origin—beginning – beginning

Successful companies weren’t always prosperous immediately. Many popular corporations today have a noteworthy story about their foundation. Some companies had an origin involving conflict when they split from an existing company. Others began with a very modest beginning, like Mark Zuckerberg beginning Facebook in his dorm room. We can better understand the present success of a company by looking at its beginning several years ago.

Int. Passage 3(D): beginning – origin—foundation – beginning

Successful companies weren’t always prosperous immediately. Many popular corporations today have a noteworthy story about their beginning. Some companies had an origin involving conflict when they split from an existing company. Others began with a very modest foundation, like Mark Zuckerberg beginning Facebook in his dorm room. We can better understand the present success of a company by looking at its beginning several years ago.

Int. Passage 3(A): beginning – beginning — beginning – beginning

Successful companies weren’t always prosperous immediately. Many popular corporations today have a noteworthy story about their beginning. Some companies had an beginning involving conflict when they split from an existing company. Others began with a very modest beginning, like Mark Zuckerberg beginning Facebook in his dorm room. We can
better understand the present success of a company by looking at its **beginning** several years ago.

**Int. Passage 3 (No Target): beginning – beginning — beginning – beginning** Successful companies weren’t always this way. Every company has an interesting story about its. Some companies had a typical by splitting from an existing company. Others began with a very modest, like Mark Zuckerberg beginning Facebook in his dorm room. Looking at a company’s helps understand its present success.

**Int. Passage 4 (N): potential – prospect—capability – possibility**

Every year, about 100 million Americans play the lottery, hoping to win big money. They play each year hoping for the **potential** to become a millionaire fast. However, the **prospect** of winning is extremely low for most popular lottery games. Generally, the **capability** is lower than the odds of getting struck by lightening. Many people play just for the **possibility** despite the low odds.

WC: 64

**True or False:** You are more likely to get hit by lightening than win the lottery.

**True or False:** This passage is about how easy it is to win the lottery.

**Int. Passage 4 (C): potential – prospect—** **possibility – possibility**

Every year, about 100 million Americans play the lottery, hoping to win big money. They play each year hoping for the **potential** to become a millionaire fast. However, the **prospect** of winning is extremely low for most popular lottery games. Generally, the
**possibility** is lower than the odds of getting struck by lightening. Many people play just for the **possibility** despite the low odds.

Int. Passage 4 (D): **possibility** – **prospect**— **potential** – **possibility**

Every year, about 100 million Americans play the lottery, hoping to win big money. They play each year hoping for the **possibility** to become a millionaire fast. However, the **prospect** of winning is extremely low for most popular lottery games. Generally, the **potential** is lower than the odds of getting struck by lightening. Many people play just for the **possibility** despite the low odds.

Int. Passage 4 (A): **possibility** – **possibility** — **possibility** – **possibility**

Every year, about 100 million Americans play the lottery, hoping to win big money. They play each year hoping for the **possibility** to become a millionaire fast. However, the **possibility** of winning is extremely low for most popular lottery games. Generally, the **possibility** is lower than the odds of getting struck by lightening. Many people play just for the **possibility** despite the low odds.

Int. Passage 4 (No Target): **possibility** – **possibility** — **possibility** – **possibility**

Every year, about 100 million Americans play the lottery, hoping to win big money. They play each year hoping for the **possibility** to win big money. However, the **possibility** of winning is extremely low. For most lottery games, the is lower than the odds of getting struck by lightening. Despite the low winning odds, many people play just for the of a lot of money.
Int. Passage 5(N): evaluation – total — mark – score

People consider some Olympic sports to be more unfair than others. In sports like figure skating, athletes have an evaluation that is subjective. Based on diverse judges’ opinions, they get a total to see who wins. This is unlike the mark from sports like soccer where you earn points. Depending on which judge is selected, athletes might receive a different score for the same performance.

WC:65

True or False: This passage is about points in Olympic sports.

True or False: The winner in soccer is decided subjectively.

Int. Passage 5(C): evaluation – total — score – score

People consider some Olympic sports to be more unfair than others. In sports like figure skating, athletes have an evaluation that is subjective. Based on diverse judges’ opinions, they get a total to see who wins. This is unlike the score from sports like soccer where you earn points. Depending on which judge is selected, athletes might receive a different score for the same performance.

Int. Passage 5(D): score – total — evaluation – score

People consider some Olympic sports to be more unfair than others. In sports like figure skating, athletes have a score that is subjective. Based on diverse judges’ opinions, they get a total to see who wins. This is unlike the evaluation from sports like soccer where you earn points. Depending on which judge is selected, athletes might receive a different score for the same performance.
Int. Passage 5(A): **score – score — score – score**

People consider some Olympic sports to be more unfair than others. In sports like figure skating, athletes have a **score** that is subjective. Based on diverse judges’ opinions, they get a **score** to see who wins. This is unlike the **score** from sports like soccer where you earn points. Depending on which judge is selected, athletes might receive a different **score** for the same performance.

Int. Passage 5(No Target): **score – score — score — score**

People consider some Olympic sports to be unfair. In sports like gymnastics or figure skating, athletes have an **that is subjective. Based on different judges’ opinions, they get a to see who wins. This is a different type of from objective sports like soccer or basketball where you score points. Depending on which judge you get, you might receive a different for the same performance.

Int. Passage 6 (N) : **proof – confirmation — testimony – evidence**

Beyonce's latest deal with Pepsi is earning her $55 million per year. Her contract states she will appear in commercials as **proof** for consumers of Pepsi's quality. This strategy of using celebrity **confirmation** is to encourage people to buy the product. Research has shown that **testimony** from celebrities raises sales. For many people, Beyonce's commercial will be enough **evidence** to buy more Pepsi.

WC: 63

**True or False:** This passage is about celebrities in advertising.

**True or False:** Beyonce is paid $100 million per year for her Pepsi advertisements.
Beyonce's latest deal with Pepsi is earning her $55 million per year. Her contract states she will appear in commercials as **proof** for consumers of Pepsi's quality. This strategy of using celebrity **confirmation** is to encourage people to buy the product. Research has shown that **evidence** from celebrities raises sales. For many people, Beyonce's commercial will be enough **evidence** to buy more Pepsi.

Int. Passage 6 (D) : **evidence** –**confirmation** — **proof** – **evidence**

Beyonce's latest deal with Pepsi is earning her $55 million per year. Her contract states she will appear in commercials as **evidence** for consumers of Pepsi's quality. This strategy of using celebrity **confirmation** is to encourage people to buy the product. Research has shown that **proof** from celebrities raises sales. For many people, Beyonce's commercial will be enough **evidence** to buy more Pepsi.

Int. Passage 6 (A) : **evidence** –**evidence** — **evidence** – **evidence**

Beyonce's latest deal with Pepsi is earning her $55 million per year. Her contract states she will appear in commercials as **evidence** for consumers of Pepsi's quality. This strategy of using celebrity **evidence** is to encourage people to buy the product. Research has shown that **evidence** from celebrities raises sales. For many people, Beyonce's commercial will be enough **evidence** to buy more Pepsi.

Int. Passage 6 (No Target): **proof** –**confirmation** — **testimony** – **evidence**

Beyonce's latest deal with Pepsi is earning her $55 million per year. Her contract states she will appear in commercials as for consumers of Pepsi's quality. This strategy of using
celebrity is to encourage people to buy the product. Research has shown that from celebrities raises sales. For many people, Beyonce's commercial will be enough to buy more Pepsi.

Int. Passage 7(N) : phase –segment —chapter – period

The U.S. has a long history of immigration. The phase where immigration was the highest was the years 1880-1930. This was an interesting segment of history in the U.S. The chapter was a time of economic heights right before the first World War and the Great Depression. After the first World War, fear of immigrants and little money changed the period quickly and dramatically.

WC: 64

True or False: This passage is about immigration in the past.

True or False: The most immigrants in the history of the U.S. came after 1930.

Int. Passage 7(C) : phase –segment — period – period

The U.S. has a long history of immigration. The phase where immigration was the highest was the years 1880-1930. This was an interesting segment of history in the U.S. The period was a time of economic heights right before the first World War and the Great Depression. After the first World War, fear of immigrants and little money changed the period quickly and dramatically.

Int. Passage 7(D) : period –segment — phase – period

The U.S. has a long history of immigration. The period where immigration was the highest was the years 1880-1930. This was an interesting segment of history in the U.S.
The phase was a time of economic heights right before the first World War and the Great Depression. After the first World War, fear of immigrants and little money changed the period quickly and dramatically.

Int. Passage 7(A) : period — period — period — period

The U.S. has a long history of immigration. The period where immigration was the highest was the years 1880-1930. This was an interesting period of history in the U.S. The period was a time of economic heights right before the first World War and the Great Depression. After the first World War, fear of immigrants and little money changed the period quickly and dramatically.

Int. Passage 7(No Target) : phase — segment — chapter — period

The U.S. has a long history of immigration. The where immigration was the highest was the years 1880-1930. This was an interesting of history in the U.S. The was a time of economic heights right before the first World War and the Great Depression. After the first World War, fear of immigrants and little money changed the quickly and dramatically.

Int. Passage 8(A) : profession — employment — occupation — career

Consider tipping your barista at Starbucks the next time you buy a mocha. Financial News Weekly just named barista one of the lowest paying professions. This employment earns $19,000 per year for a full-time employee working 40-hour weeks. That means that this occupation gives you just $4,000 more than the poverty level. Despite the low pay, many baristas say they love their career even without high pay.
True or False: This passage is about how Starbucks employees don't make much money.

True or False: Baristas earn less than the poverty level amount of money.

Int. Passage 8(C) : profession – employment – career – career

Consider tipping your barista at Starbucks the next time you buy a mocha. Financial News Weekly just named barista one of the lowest paying professions. This employment earns $19,000 per year for a full-time employee working 40-hour weeks. That means that this career gives you just $4,000 more than the poverty level. Despite the low pay, many baristas say they love their career even without high pay.

Int. Passage 8(D) : career – employment – profession – career

Consider tipping your barista at Starbucks the next time you buy a mocha. Financial News Weekly just named barista one of the lowest paying careers. This employment earns $19,000 per year for a full-time employee working 40-hour weeks. That means that this profession gives you just $4,000 more than the poverty level. Despite the low pay, many baristas say they love their career even without high pay.

Int. Passage 8(A) : career – career – career – career

Consider tipping your barista at Starbucks the next time you buy a mocha. Financial News Weekly just named barista one of the lowest paying careers. This career earns $19,000 per year for a full-time employee working 40-hour weeks. That means that this career gives you just $4,000 more than the poverty level. Despite the low pay, many baristas say they love their career even without high pay.
Consider tipping your barista at Starbucks the next time you buy a mocha. Financial News Weekly just named barista one of the lowest paying. This earns $19,000 per year for a full-time employee working 40-hour weeks. That means that this gives you just $4,000 more than the poverty level. Despite the low pay, many baristas say they love their even without high pay.

Climate change has been a controversial political topic recently. Scientist Michael Mann developed a chart of global temperatures since 1400 AD. The drawing has been both praised and criticized by many people. People don't like to believe the temperatures are rising as indicated by the representation. However, scientists continue to educate the people using this figure to stop pollution.

True or False: This passage is disagreement over climate change.

True or False: Michael Mann reports on climate change.

Climate change has been a controversial political topic recently. Scientist Michael Mann developed a chart of global temperatures since 1400 AD. The drawing has been both praised and criticized by many people. People don't like to believe the temperatures are rising as indicated by the figure. However, scientists continue to educate the people using this figure to stop pollution.
Int. Passage 9(D) : **figure** – **drawing** — **chart** – **figure**

Climate change has been a controversial political topic recently. Scientist Michael Mann developed a **figure** of global temperatures since 1400 AD. The **drawing** has been both praised and criticized by many people. People don't like to believe the temperatures are rising as indicated by the **chart**. However, scientists continue to educate the people using this **figure** to stop pollution.

Int. Passage 9(A) : **figure** – **figure** — **figure** – **figure**

Climate change has been a controversial political topic recently. Scientist Michael Mann developed a **figure** of global temperatures since 1400 AD. The **figure** has been both praised and criticized by many people. People don't like to believe the temperatures are rising as indicated by the **figure**. However, scientists continue to educate the people using this **figure** to stop pollution.

Int. Passage 9( No Target): **chart** – **drawing** — **representation** – **figure**

Climate change has been a controversial political topic recently. Scientist Michael Mann developed a of global temperatures since 1400 AD. The **has been both praised and criticized by many people. People don't like to believe the temperatures are rising as indicated by the. However, scientists continue to educate the people using this to stop pollution.

Int. Passage 10(N) : **diversity** – **variation** — **distinction** – **difference**

In the 1950s, more than 60% of the workers in America were white men. Over the last 50 years, there has been an increase in **diversity** among workers. The **variation** means there
are more perspectives and thus different ideas or solutions. With greater **distinction**, there are also more talents being used in companies. The majority of companies view the **differences** as something positive.

**WC: 63**

**True** or False: This passage is about employees in the U.S.

**True** or False: Companies think the changes over the last 50 years in employees are positive.

**Int. Passage 10(C) :** diversity – variation — **difference** – **difference**

In the 1950s, more than 60% of the workers in America were white men. Over the last 50 years, there has been an increase in **diversity** among workers. The **variation** means there are more perspectives and thus different ideas or solutions. With greater **differences**, there are also more talents being used in companies. The majority of companies view the **differences** as something positive.

**Int. Passage 10(D) :** **difference** – variation — diversity – **difference**

In the 1950s, more than 60% of the workers in America were white men. Over the last 50 years, there has been an increase in **differences** among workers. The **variation** means there are more perspectives and thus different ideas or solutions. With greater **diversity**, there are also more talents being used in companies. The majority of companies view the **differences** as something positive.
Int. Passage 10(A): **difference – difference — difference – difference**

In the 1950s, more than 60% of the workers in America were white men. Over the last 50 years, there has been an increase in differences among workers. The differences mean there are more perspectives and thus different ideas or solutions. With greater differences, there are also more talents being used in companies. The majority of companies view the differences as something positive.

Int. Passage 10(No Targets): **diversity – variation — distinction — difference**

In the 1950s, more than 60% of the workers in America were white men. Over the last 50 years, there has been an increase in among workers. The means there are more perspectives and thus different ideas or solutions. With greater, there are also more talents being used in companies. The majority of companies view the as something positive.

Int. Passage 11(N): **ranks – promotion — standing – status**

Most young soldiers entering the U.S. army begin by being called a *Private*. Actually, there are over thirteen different ranks that people can achieve during their military service. These standings indicate demonstration of leadership and commitment to the army. There are many serious responsibilities that come with the promotions. For example, making decisions is dependent on your status in the army.

WC: 61

**True or False:** This passage is about positions in the U.S. military.

**True or False:** Not everyone makes decisions in the U.S. army.
Most young soldiers entering the U.S. army begin by being called a *Private*. Actually, there are over thirteen different ranks that people can achieve during their military service. These standings indicate demonstration of leadership and commitment to the army. There are many serious responsibilities that come with the status. For example, making decisions is dependent on your status in the army.

Most young soldiers entering the U.S. army begin by being called a *Private*. Actually, there are over thirteen different status that people can achieve during their military service. These standings indicate demonstration of leadership and commitment to the army. There are many serious responsibilities that come with the ranks. For example, making decisions is dependent on your status in the army.

Most young soldiers entering the U.S. army begin by being called a *Private*. Actually, there are over thirteen different status that people can achieve during their military service. These status indicate demonstration of leadership and commitment to the army. There are many serious responsibilities that come with the status. For example, making decisions is dependent on your status in the army.
These indicate demonstration of leadership and commitment to the army. There are many serious responsibilities that come with the. For example, making decisions is dependent on your in the army.

Int. Passage 12(N): adjustment — adaptation — alteration — change

Peter the Great was very influential in shaping the domestic structure of Russia. One of his significant adjustments was to military education in his country. After travelling, he made the adaptation because he realized other militaries had educated soldiers. Part of the alteration was establishing different schools for engineering and mathematics. The public of Russia considered the change to be beneficial.

WC: 61

True or False: This passage is about Peter the Great.

True or False: Peter the Great established multiple schools in Russia.

Int. Passage 12(C): adjustment — adaptation — change — change

Peter the Great was very influential in shaping the domestic structure of Russia. One of his significant adjustments was to military education in his country. After travelling, he made the adaptation because he realized other militaries had educated soldiers. Part of the change was establishing different schools for engineering and mathematics. The public of Russia considered the change to be beneficial.

Int. Passage 12(D): change — adaptation — adjustment — change

Peter the Great was very influential in shaping the domestic structure of Russia. One of his significant changes was to military education in his country. After travelling, he made
the adaptation because he realized other militaries had educated soldiers. Part of the adjustment was establishing different schools for engineering and mathematics. The public of Russia considered the change to be beneficial.

Int. Passage 12(A): change – change — change – change

Peter the Great was very influential in shaping the domestic structure of Russia. One of his significant changes was to military education in his country. After travelling, he made the change because he realized other militaries had educated soldiers. Part of the change was establishing different schools for engineering and mathematics. The public of Russia considered the change to be beneficial.

Int. Passage 12(No Targets): adjustment – adaptation — alteration — change

Peter the Great was very influential in shaping the domestic structure of Russia. One of his significant was to military education in his country. After travelling, he made the because he realized other militaries had educated soldiers. Part of the was establishing different schools for engineering and mathematics. The public of Russia considered the to be beneficial.

Int. Passage 13(N): headquarters — complex — campus — facility

The Coca Cola Company was founded in Atlanta, Georgia. The company headquarters was established in 1979 by designer Bob Foker. The complex includes a 29-floor tower and a large coca cola structure in the front. Nearly 300,000 tourists visit the campus each year to take a tour and learn more about their favorite beverage. No trip to Georgia is complete without visiting facility for a few hours.
True or False: This passage is about how to make Coca Cola.

True or False: Coca Cola has a 10-story building.

Int. Passage 13(C): headquarters – complex — facility – facility

The Coca Cola Company was founded in Atlanta, Georgia. The company headquarters was established in 1979 by designer Bob Foker. The complex includes a 29-floor tower and a large coca cola structure in the front. Nearly 300,000 tourists visit the facility each year to take a tour and learn more about their favorite beverage. No trip to Georgia is complete without visiting facility for a few hours.

Int. Passage 13(D): facility – complex — headquarters – facility

The Coca Cola Company was founded in Atlanta, Georgia. The company facility was established in 1979 by designer Bob Foker. The complex includes a 29-floor tower and a large coca cola structure in the front. Nearly 300,000 tourists visit the headquarters each year to take a tour and learn more about their favorite beverage. No trip to Georgia is complete without visiting facility for a few hours.

Int. Passage 13(A): facility – facility — facility – facility

The Coca Cola Company was founded in Atlanta, Georgia. The company facility was established in 1979 by designer Bob Foker. The facility includes a 29-floor tower and a large coca cola structure in the front. Nearly 300,000 tourists visit the facility each year to take a tour and learn more about their favorite beverage. No trip to Georgia is complete without visiting facility for a few hours.
Int. Passage 13(No Target): headquarters –complex —campus – facility

The Coca Cola Company was founded in Atlanta, Georgia. The company was established in 1979 by designer Bob Foker. The includes a 29-floor tower and a large coca cola structure in the front. Nearly 300,000 tourists visit the each year to take a tour and learn more about their favorite beverage. No trip to Georgia is complete without visiting for a few hours.

Int. Passage 14(N): adventure –journey —exploration – experience

The Appalachian trail, stretching from Georgia to Maine, is famous among hikers. The first man to ever have the adventure of hiking the Appalachian Trail is Earl Shaffer. In 1948, he started his journey after he finished fighting in World War II. It took him 124 days to complete the exploration, walking about 16.5 miles per day. Shaffer had no regrets and said the experience was the best.

WC: 68

True or False: This passage is about what to bring with you when hiking the Appalachian Trail.

True or False: Earl Shaffer never finished hiking the Appalachian Trail.

Int. Passage 14(C): adventure –journey — experience – experience

The Appalachian trail, stretching from Georgia to Maine, is famous among hikers. The first man to ever have the adventure of hiking the Appalachian Trail is Earl Shaffer. In 1948, he started his journey after he finished fighting in World War II. It took him 124
days to complete the experience, walking about 16.5 miles per day. Shaffer had no regrets and said the experience was the best.

Int. Passage 14(D): experience – journey — adventure – experience

The Appalachian trail, stretching from Georgia to Maine, is famous among hikers. The first man to ever have the experience of hiking the Appalachian Trail is Earl Shaffer. In 1948, he started his journey after he finished fighting in World War II. It took him 124 days to complete the adventure, walking about 16.5 miles per day. Shaffer had no regrets and said the experience was the best.

Int. Passage 14(A): experience – experience — experience – experience

The Appalachian trail, stretching from Georgia to Maine, is famous among hikers. The first man to ever have the experience of hiking the Appalachian Trail is Earl Shaffer. In 1948, he started his experience after he finished fighting in World War II. It took him 124 days to complete the experience, walking about 16.5 miles per day. Shaffer had no regrets and said the experience was the best.

Int. Passage 14(No Targets): adventure – journey — exploration — experience

The Appalachian trail, stretching from Georgia to Maine, is famous among hikers. The first man to ever have the of hiking the Appalachian Trail is Earl Shaffer. In 1948, he started his after he finished fighting in World War II. It took him 124 days to complete the, walking about 16.5 miles per day. Shaffer had no regrets and said the was the best.
Ombudsmen are people who facilitate problems between high-power people and low-power people. For example, they need the wisdom to know how to work with both a CEO and a temporary worker. Typically, these people have a lot of insight about people and relationships. Because they deal with many problems, their capability helps them to be fair to everyone. Not everyone has the knowledge to be an ombudsman.

WC:67

True or False: This passage is about how CEOs work with their employees.

True or False: Ombudsmen don't have any special skills.
Int. Passage 15(A): **knowledge — knowledge — knowledge — knowledge**

Ombudsmen are people who facilitate problems between high-power people and low-power people. For example, they need the **knowledge** to know how to work with both a CEO and a temporary worker. Typically, these people have a lot of **knowledge** about people and relationships. Because they deal with many problems, their **knowledge** helps them to be fair to everyone. Not everyone has the **knowledge** to be an ombudsman.

Int. Passage 15(No Target): **wisdom — insight — capability — knowledge**

Ombudsmen are people who facilitate problems between high-power people and low-power people. For example, they need the to know how to work with both a CEO and a temporary worker. Typically, these people have a lot of about people and relationships. Because they deal with many problems, their helps them to be fair to everyone. Not everyone has the to be an ombudsman.

Int. Passage 16 (N): **selection — mixture — arrangement — variety**

Macy's Department store first started as a grocery store in 1858. Now, it sells a wide **selection** of goods and products to modern customers. The **mixture** appeals to a wide variety of Americans as their stores are in 45 states. The **arrangement** includes things like clothing, appliances, and even technology. The store has advanced tremendously thanks to the **variety** that it has.

WC: 62

True or **False**: This passage is about the history of grocery stores.

True or **False**: Macy's is located in all 50 states.
Macy’s Department store first started as a grocery store in 1858. Now, it sells a wide selection of goods and products to modern customers. The mixture appeals to a wide variety of Americans as their stores are in 45 states. The variety includes things like clothing, appliances, and even technology. The store has advanced tremendously thanks to the variety that it has.

Macy’s Department store first started as a grocery store in 1858. Now, it sells a wide variety of goods and products to modern customers. The mixture appeals to a wide variety of Americans as their stores are in 45 states. The selection includes things like clothing, appliances, and even technology. The store has advanced tremendously thanks to the variety that it has.

Macy’s Department store first started as a grocery store in 1858. Now, it sells a wide variety of goods and products to modern customers. The variety appeals to a wide variety of Americans as their stores are in 45 states. The variety includes things like clothing, appliances, and even technology. The store has advanced tremendously thanks to the variety that it has.

Macy's Department store first started as a grocery store in 1858. Now, it sells a wide of goods and products to modern customers. The appeals to a wide variety of Americans as
their stores are in 45 states. The includes things like clothing, appliances, and even technology. The store has advanced tremendously thanks to the that it has.

**Advanced Passages**

Adv. Passage 1(N) : obligation -- necessity -- constraint -- requirement

Individual states are adjusting laws to reduce teenage driving accidents. New Jersey is leading the way with obligations on drivers under 21. They instituted several license necessities, such as imposing a curfew. Police ensure teens abide by these constraints by the red license plate sticker on underage drivers’ cars. In the past two years, accidents have diminished due to the requirements on young drivers.

WC: 64

Adv. Passage 1(C) : obligation -- necessity -- requirement -- requirement

Individual states are adjusting laws to reduce teenage driving accidents. New Jersey is leading the way with obligations on drivers under 21. They instituted several license necessities, such as imposing a curfew. Police ensure teens abide by these requirements by the red license plate sticker on underage drivers’ cars. In the past two years, accidents have diminished due to the requirements on young drivers.

Adv. Passage 1(D) : requirement -- necessity -- obligation -- requirement

Individual states are adjusting laws to reduce teenage driving accidents. New Jersey is leading the way with requirements on drivers under 21. They instituted several license necessities, such as imposing a curfew. Police ensure teens abide by these obligations by
the red license plate sticker on underage drivers’ cars. In the past two years, accidents have diminished due to the requirements on young drivers.

Adv. Passage 1(A) : requirement -- requirement -- requirement -- requirement

Individual states are adjusting laws to reduce teenage driving accidents. New Jersey is leading the way with requirements on drivers under 21. They instituted several license requirements, such as imposing a curfew. Police ensure teens abide by these requirements by the red license plate sticker on underage drivers’ cars. In the past two years, accidents have diminished due to the requirements on young drivers.

Adv. Passage 1 No Targets

Individual states are adjusting laws to reduce teenage driving accidents. New Jersey is leading the way with on drivers under 21. They instituted several license, such as imposing a curfew. Police ensure teens abide by these by the red license plate sticker on underage drivers’ cars. In the past two years, accidents have diminished due to the on young drivers.


Marriage regulations are very strict in the United States. People younger than 18 years old must have consent from their guardians. It is necessary to get authorization from the parents because teenagers are susceptible to making poor decisions. Without permission, young people cannot obtain the necessary license. Under these laws, there is no way to avoid having agreement from adults before marriage.

WC: 62
Marriage regulations are very strict in the United States. People younger than 18 years old must have consent from their guardians. It is necessary to get authorization from the parents because teenagers are susceptible to making poor decisions. Without agreement, young people cannot obtain the necessary license. Under these laws, there is no way to avoid having agreement from adults before marriage.

Marriage regulations are very strict in the United States. People younger than 18 years old must have agreement from their guardians. It is necessary to get agreement from the parents because teenagers are susceptible to making poor decisions. Without consent, young people cannot obtain the necessary license. Under these laws, there is no way to avoid having agreement from adults before marriage.

Marriage regulations are very strict in the United States. People younger than 18 years old must have agreement from their guardians. It is necessary to get agreement from the parents because teenagers are susceptible to making poor decisions. Without agreement, young people cannot obtain the necessary license. Under these laws, there is no way to avoid having agreement from adults before marriage.

Marriage regulations are very strict in the United States. People younger than 18 years old must have agreement from their guardians. It is necessary to get agreement from the parents because teenagers are susceptible to making poor decisions. Without agreement, young people cannot obtain the necessary license. Under these laws, there is no way to avoid having agreement from adults before marriage.
teenagers are susceptible to making poor decisions. Without, young people cannot obtain the necessary license. Under these laws, there is no way to avoid having from adults before marriage.


Starting a small business is no easy feat to accomplish these days. There are a number of obstacles to success in entrepreneurship. One of the first complications involves finding a bank to support your loan. Hiring trustworthy employees and equipping them well are other hurdles to consider. Owning your own business is very rewarding once the initial difficulties can be overcome.

WC: 61

Adv. Passage 3(C): obstacle – complication – difficulty —difficulty

Starting a small business is no easy feat to accomplish these days. There are a number of obstacles to success in entrepreneurship. One of the first complications involves finding a bank to support your loan. Hiring trustworthy employees and equipping them well are other difficulties to consider. Owning your own business is very rewarding once the initial difficulties can be overcome.

Adv. Passage 3(D): obstacle – complication – difficulty —difficulty

Starting a small business is no easy feat to accomplish these days. There are a number of difficulties to success in entrepreneurship. One of the first complications involves finding a bank to support your loan. Hiring trustworthy employees and equipping them
well are other **obstacles** to consider. Owning your own business is very rewarding once the initial **difficulties** can be overcome.

Adv. Passage 3(A): **difficulty** –  **difficulty**  –  **difficulty**  — **difficulty**

Starting a small business is no easy feat to accomplish these days. There are a number of **difficulties** to success in entrepreneurship. One of the first **difficulties** involves finding a bank to support your loan. Hiring trustworthy employees and equipping them well are other **difficulties** to consider. Owning your own business is very rewarding once the initial **difficulties** can be overcome.

Adv. Passage 3: No targets

Starting a small business is no easy feat to accomplish these days. There are a number of **difficulties** to success in entrepreneurship. One of the first involves finding a bank to support your loan. Hiring trustworthy employees and equipping them well are other to consider. Owning your own business is very rewarding once the initial can be overcome.

Adv. Passage 4(N):  **panic** –  **worry**—**nervousness** –  **fear**

Becoming a parent for the first time is both rewarding and overwhelming. Initially, parents typically experience **panic** when bringing the infant home from the hospital. New mothers and fathers can’t always discern the babies’ needs and induces their **nervousness**. Over time, they gain experience and wisdom, and their **worry** decreases.

Before the child is one year old, the parents have less **fear** and feel better about parenthood.

WC: 67

Becoming a parent for the first time is both rewarding and overwhelming. Initially, parents typically experience panic when bringing the infant home from the hospital. New mothers and fathers can’t always discern the babies’ needs and induces their nervousness. Over time, they gain experience and wisdom, and their fear decreases. Before the child is one year old, the parents have less fear and feel better about parenthood.


Becoming a parent for the first time is both rewarding and overwhelming. Initially, parents typically experience fear when bringing the infant home from the hospital. New mothers and fathers can’t always discern the babies’ needs and induces their nervousness. Over time, they gain experience and wisdom, and their panic decreases. Before the child is one year old, the parents have less fear and feel better about parenthood.


Becoming a parent for the first time is both rewarding and overwhelming. Initially, parents typically experience fear when bringing the infant home from the hospital. New mothers and fathers can’t always discern the babies’ needs and induces their fear. Over time, they gain experience and wisdom, and their fear decreases. Before the child is one year old, the parents have less fear and feel better about parenthood.
Adv. Passage 4 No Targets

Becoming a parent for the first time is both rewarding and overwhelming. Initially, parents typically experience when bringing the infant home from the hospital. New mothers and fathers can’t always discern the babies’ needs and induces their. Over time, they gain experience and wisdom, and their decreases. Before the child is one year old, the parents have less and feel better about parenthood.


The best employees are the ones who are screened carefully in the interview. Employees displaying a feeling of excitement for their job tend to work harder. Their delight enables them to do good work with a positive attitude. Employees having a lot of enthusiasm will make any company more successful. When conducting interviews, always see how much pleasure they have before hiring anyone.

WC: 63

Adv. Passage 5(C): excitement – delight —pleasure— pleasure

The best employees are the ones who are screened carefully in the interview. Employees displaying a feeling of excitement for their job tend to work harder. Their delight enables them to do good work with a positive attitude. Employees having a lot of pleasure will make any company more successful. When conducting interviews, always see how much pleasure they have before hiring anyone.
Adv. Passage 5(D): **pleasure** – delight — excitement — pleasure

The best employees are the ones who are screened carefully in the interview. Employees displaying a feeling of pleasure for their job tend to work harder. Their delight enables them to do good work with a positive attitude. Employees having a lot of excitement will make any company more successful. When conducting interviews, always see how much pleasure they have before hiring anyone.

Adv. Passage 5(A): **pleasure** – pleasure—pleasure—pleasure

The best employees are the ones who are screened carefully in the interview. Employees displaying a feeling of pleasure for their job tend to work harder. Their pleasure enables them to do good work with a positive attitude. Employees having a lot of pleasure will make any company more successful. When conducting interviews, always see how much pleasure they have before hiring anyone.

Adv. Passage 5 No Targets

The best employees are the ones who are screened carefully in the interview. Employees displaying a feeling of for their job tend to work harder. Their enables them to do good work with a positive attitude. Employees having a lot of will make any company more successful. When conducting interviews, always see how much they have before hiring anyone.


Living with a roommate in college is not an easy feat. Plenty of students have never had a communal room before, so disagreements are typical. Most roommates have clashes
over how to keep the room tidy. Although *confrontations* are never pleasant, they teach students how to communicate. Sharing a small space will prepare them for the *fights* they will have later in life.

Adv. Passage 6(C): disagreement – clash – *fight* – *fight*

Living with a roommate in college is not an easy feat. Plenty of students have never had a communal room before, so *disagreements* are typical. Most roommates have *clashes* over how to keep the room tidy. Although *fights* are never pleasant, they teach students how to communicate. Sharing a small space will prepare them for the *fights* they will have later in life.

Adv. Passage 6(D): *fight* – clash – disagreement – *fight*

Living with a roommate in college is not an easy feat. Plenty of students have never had a communal room before, so *fights* are typical. Most roommates have *clashes* over how to keep the room tidy. Although *disagreements* are never pleasant, they teach students how to communicate. Sharing a small space will prepare them for the *fights* they will have later in life.

Adv. Passage 6(A): *fight* – *fight* – *fight* – *fight*

Living with a roommate in college is not an easy feat. Plenty of students have never had a communal room before, so *fights* are typical. Most roommates have *fights* over how to keep the room tidy. Although *fights* are never pleasant, they teach students how to
communicate. Sharing a small space will prepare them for the fights they will have later in life.

**Adv. Passage 6 No Targets**

Living with a roommate in college is not an easy feat. Plenty of students have never had a communal room before, so are typical. Most roommates have over how to keep the room tidy. Although are never pleasant, they teach students how to communicate. Sharing a small space will prepare them for the fights they will have later in life.


Business students are generally the most adventurous type of undergraduate. They are more prone to take on a significant venture while in college. For some, this enterprise is founding a club or organizing a student movement. For others, their pursuit is getting a prestigious internship with a firm they’d like to work for. Students are usually grateful they took on such a mission after they graduate.

**WC:66**


Business students are generally the most adventurous type of undergraduate. They are more prone to take on a significant venture while in college. For some, this enterprise is founding a club or organizing a student movement. For others, their mission is getting a prestigious internship with a firm they’d like to work for. Students are usually grateful they took on such a mission after they graduate.

**Adv. Passage 7 (D): mission – enterprise—venture – mission**
Business students are generally the most adventurous type of undergraduate. They are more prone to take on a significant mission while in college. For some, this enterprise is founding a club or organizing a student movement. For others, their venture is getting a prestigious internship with a firm they’d like to work for. Students are usually grateful they took on such a mission after they graduate.


Business students are generally the most adventurous type of undergraduate. They are more prone to take on a significant mission while in college. For some, this mission is founding a club or organizing a student movement. For others, their mission is getting a prestigious internship with a firm they’d like to work for. Students are usually grateful they took on such a mission after they graduate.

Adv. Passage 7 No Targets

Business students are generally the most adventurous type of undergraduate. They are more prone to take on a significant while in college. For some, this is founding a club or organizing a student movement. For others, their is getting a prestigious internship with a firm they’d like to work for. Students are usually grateful they took on such a after they graduate.

Adv. Passage 8(N): outline — summary—sketch — idea

In 1995, Jeff Bezos quit his job and moved to Seattle to begin Amazon. He first shared his outline for an online book distribution site with his friends. He and his friends discussed the summary in his garage and at nearby bookstores. In just two years, his
sketch turned into a very successful online business. Now, many other online retailers are copying his idea to be successful.

Adv. Passage 8(C): outline – summary—idea – idea

In 1995, Jeff Bezos quit his job and moved to Seattle to begin Amazon. He first shared his outline for an online book distribution site with his friends. He and his friends discussed the summary in his garage and at nearby bookstores. In just two years, his idea turned into a very successful online business. Now, many other online retailers are copying his idea to be successful.

Adv. Passage 8(D): idea – summary—outline – idea

In 1995, Jeff Bezos quit his job and moved to Seattle to begin Amazon. He first shared his idea for an online book distribution site with his friends. He and his friends discussed the summary in his garage and at nearby bookstores. In just two years, his outline turned into a very successful online business. Now, many other online retailers are copying his idea to be successful.


In 1995, Jeff Bezos quit his job and moved to Seattle to begin Amazon. He first shared his idea for an online book distribution site with his friends. He and his friends discussed the idea in his garage and at nearby bookstores. In just two years, his idea turned into a very successful online business. Now, many other online retailers are copying his idea to be successful.
Adv. Passage 8 No Targets

In 1995, Jeff Bezos quit his job and moved to Seattle to begin Amazon. He first shared his idea for an online book distribution site with his friends. He and his friends discussed the idea in his garage and at nearby bookstores. In just two years, his idea turned into a very successful online business. Now, many other online retailers are copying his idea to be successful.


The pollution of the Earth's oceans is growing increasingly worse. Thankfully, Australian surfers have come up with an invention to help curb the problem. Their innovation is a floating trash can with a pump that circulates water through, collecting trash. Called the "Seabin," this breakthrough will be made of recyclable materials. With nearly $300,000 already raised, it won't be long before this device is in the oceans.

WC: 67

Adv. Passage 9(C): invention – innovation—device – device

The pollution of the Earth's oceans is growing increasingly worse. Thankfully, Australian surfers have come up with an invention to help curb the problem. Their innovation is a floating trash can with a pump that circulates water through, collecting trash. Called the "Seabin," this device will be made of recyclable materials. With nearly $300,000 already raised, it won't be long before this device is in the oceans.

Adv. Passage 9(D): device – innovation— invention – device

The pollution of the Earth's oceans is growing increasingly worse. Thankfully, Australian surfers have come up with an device to help curb the problem. Their innovation is a
floating trash can with a pump that circulates water through, collecting trash. Called the "Seabin," this invention will be made of recyclable materials. With nearly $300,000 already raised, it won't be long before this device is in the oceans.

The pollution of the Earth's oceans is growing increasingly worse. Thankfully, Australian surfers have come up with a device to help curb the problem. Their device is a floating trash can with a pump that circulates water through, collecting trash. Called the "Seabin," this device will be made of recyclable materials. With nearly $300,000 already raised, it won't be long before this device is in the oceans.

On average, a presidential candidate in the United States will require around $200 million for their campaign. The majority of these funds come from donations from wealthy backers. Other offerings come from corporations or industries with political interests similar to the candidate. The candidate may use these subsidies for advertisements or travelling purposes. Without enough of this support, candidates drop out of the race.
Adv. Passage 10 (C): donation – offering— support – support

On average, a presidential candidate in the United States will require around $200 million for their campaign. The majority of these funds come from donations from wealthy backers. Other offerings come from corporations or industries with political interests similar to the candidate. The candidate may use these support for advertisements or travelling purposes. Without enough of this support, candidates drop out of the race.

Adv. Passage 10 (D): support – offering— donation – support

On average, a presidential candidate in the United States will require around $200 million for their campaign. The majority of these funds come from support from wealthy backers. Other offerings come from corporations or industries with political interests similar to the candidate. The candidate may use these donations for advertisements or travelling purposes. Without enough of this support, candidates drop out of the race.

Adv. Passage 10 (A): support – support — support – support

On average, a presidential candidate in the United States will require around $200 million for their campaign. The majority of these funds come from support from wealthy backers. Other support come from corporations or industries with political interests similar to the candidate. The candidate may use these support for advertisements or travelling purposes. Without enough of this support, candidates drop out of the race.
On average, a presidential candidate in the United States will require around $200 million for their campaign. The majority of these funds come from wealthy backers. Other come from corporations or industries with political interests similar to the candidate. The candidate may use these for advertisements or travelling purposes. Without enough of this, candidates drop out of the race.

Finished in May 2016, The Valravn is the world's new tallest roller coaster. This completion is celebrated by the amusement park, Cedar Point, located in Ohio. The triumph took over three years of planning, construction, and testing. Cedar Point's accomplishment knocks Kingda Ka in New Jersey from the top spot. More visitors will undoubtedly flock to Ohio to see the 500-ft tall victory in real life.
Adv. Passage 11(D): **victory** – accomplishment— **triumph** – **victory**

Finished in May 2016, The Valravn is the world's new tallest roller coaster. This **victory** is celebrated by the amusement park, Cedar Point, located in Ohio. The **triumph** took over three years of planning, construction, and testing. Cedar Point's **accomplishment** knocks Kingda Ka in New Jersey from the top spot. More visitors will undoubtedly flock to Ohio to see the 500-ft tall **victory** in real life.

Adv. Passage 11(A): **victory** – **victory** — **victory** – **victory**

Finished in May 2016, The Valravn is the world's new tallest roller coaster. This **victory** is celebrated by the amusement park, Cedar Point, located in Ohio. The **victory** took over three years of planning, construction, and testing. Cedar Point's **victory** knocks Kingda Ka in New Jersey from the top spot. More visitors will undoubtedly flock to Ohio to see the 500-ft tall **victory** in real life.

Adv. Passage 11(No Targets): **triumph** – accomplishment—completion – **victory**

Finished in May 2016, The Valravn is the world's new tallest roller coaster. This is celebrated by the amusement park, Cedar Point, located in Ohio. The **triumph** took over three years of planning, construction, and testing. Cedar Point's **victory** knocks Kingda Ka in New Jersey from the top spot. More visitors will undoubtedly flock to Ohio to see the 500-ft tall in real life.

Adv. Passage 12(N): **motive** – aim—ambition – **purpose**

Bethany Hamilton is a surfer who lost her arm in a shark attack. Wanting to win a surfing championship, she had a **motive** to recover quickly. Bethany pursued her **aim** by
beginning surfing again after just one month. Because of her ambition, she was able to overcome the traumatic event mentally and physically. As depicted in the movie Soul Surfer, she fulfilled her purpose in just two years.

Adv. Passage 12(C): motive – aim— purpose – purpose

Bethany Hamilton is a surfer who lost her arm in a shark attack. Wanting to win a surfing championship, she had a motive to recover quickly. Bethany pursued her aim by beginning surfing again after just one month. Because of her purpose, she was able to overcome the traumatic event mentally and physically. As depicted in the movie Soul Surfer, she fulfilled her purpose in just two years.


Bethany Hamilton is a surfer who lost her arm in a shark attack. Wanting to win a surfing championship, she had a purpose to recover quickly. Bethany pursued her aim by beginning surfing again after just one month. Because of her motive, she was able to overcome the traumatic event mentally and physically. As depicted in the movie Soul Surfer, she fulfilled her purpose in just two years.


Bethany Hamilton is a surfer who lost her arm in a shark attack. Wanting to win a surfing championship, she had a purpose to recover quickly. Bethany pursued her purpose by beginning surfing again after just one month. Because of her purpose, she was able to
overcome the traumatic event mentally and physically. As depicted in the movie *Soul Surfer*, she fulfilled her purpose in just two years.

**Adv. Passage 12(No Targets): motive – aim—ambition – purpose**

Bethany Hamilton is a surfer who lost her arm in a shark attack. Wanting to win a surfing championship, she had a to recover quickly. Bethany pursued her by beginning surfing again after just one month. Because of her, she was able to overcome the traumatic event mentally and physically. As depicted in the movie *Soul Surfer*, she fulfilled her in just two years.


Over the weekend, police were called to handle a violent situation where Donald Trump was giving a speech. Trump protestors were outside with signs and Trump supporters joined the gathering. Over 300 people from both sides were roaring at each other at the assembly. The rally quickly became overrun with physical fighting and violence. The police arrested over 52 participants from the crowd before people left calmly.

**WC: 67**

**Adv. Passage 13(C): gathering – assembly— crowd – crowd**

Over the weekend, police were called to handle a violent situation where Donald Trump was giving a speech. Trump protestors were outside with signs and Trump supporters joined the gathering. Over 300 people from both sides were roaring at each other at the assembly. The crowd quickly became overrun with physical fighting and violence. The police arrested over 52 participants from the crowd before people left calmly.
Over the weekend, police were called to handle a violent situation where Donald Trump was giving a speech. Trump protestors were outside with signs and Trump supporters joined the crowd. Over 300 people from both sides were roaring at each other at the assembly. The gathering quickly became overrun with physical fighting and violence. The police arrested over 52 participants from the crowd before people left calmly.

Over the weekend, police were called to handle a violent situation where Donald Trump was giving a speech. Trump protestors were outside with signs and Trump supporters joined the crowd. Over 300 people from both sides were roaring at each other at the crowd. The crowd quickly became overrun with physical fighting and violence. The police arrested over 52 participants from the crowd before people left calmly.

Over the weekend, police were called to handle a violent situation where Donald Trump was giving a speech. Trump protestors were outside with signs and Trump supporters joined the. Over 300 people from both sides were roaring at each other at the. The quickly became overrun with physical fighting and violence. The police arrested over 52 participants from the before people left calmly.

On August 9, 1974, U.S. President Richard Nixon made an announcement to the American people. He announced his departure from the office of president for the
betterment of the country. Many people were demanding his withdrawal because he had lied about a scandal. The removal made Gerald Ford the new president of the U.S. People still question the motives for his leaving to this day.

WC: 65


On August 9, 1974, U.S. President Richard Nixon made an announcement to the American people. He announced his departure from the office of president for the betterment of the country. Many people were demanding his withdrawal because he had lied about a scandal. The leaving made Gerald Ford the new president of the U.S. People still question the motives for his leaving to this day.


On August 9, 1974, U.S. President Richard Nixon made an announcement to the American people. He announced his leaving from the office of president for the betterment of the country. Many people were demanding his withdrawal because he had lied about a scandal. The departure made Gerald Ford the new president of the U.S. People still question the motives for his leaving to this day.


On August 9, 1974, U.S. President Richard Nixon made an announcement to the American people. He announced his leaving from the office of president for the betterment of the country. Many people were demanding his leaving because he had lied
about a scandal. The leaving made Gerald Ford the new president of the U.S. People still question the motives for his leaving to this day.


On August 9, 1974, U.S. President Richard Nixon made an announcement to the American people. He announced his from the office of president for the betterment of the country. Many people were demanding his because he had lied about a scandal. The made Gerald Ford the new president of the U.S. People still question the motives for his to this day.


The holy grail is mythical cup supposedly able to give eternal youth. Since the 12th century, people have led hunts to find the holy grail. Most people lead their quests to England, the last rumored location. Stories of these expeditions have been used as the plots of countless movies and books. Despite the popularity of the myth, no search has ever been successful.

WC:63


The holy grail is mythical cup supposedly able to give eternal youth. Since the 12th century, people have led hunts to find the holy grail. Most people lead their quests to England, the last rumored location. Stories of these searches have been used as the plots of countless movies and books. Despite the popularity of the myth, no search has ever been successful.
The holy grail is mythical cup supposedly able to give eternal youth. Since the 12th century, people have led searches to find the holy grail. Most people lead their quests to England, the last rumored location. Stories of these hunts have been used as the plots of countless movies and books. Despite the popularity of the myth, no search has ever been successful.

The holy grail is mythical cup supposedly able to give eternal youth. Since the 12th century, people have led searches to find the holy grail. Most people lead their searches to England, the last rumored location. Stories of these searches have been used as the plots of countless movies and books. Despite the popularity of the myth, no search has ever been successful.

The holy grail is mythical cup supposedly able to give eternal youth. Since the 12th century, people have led to find the holy grail. Most people lead their to England, the last rumored location. Stories of these have been used as the plots of countless movies and books. Despite the popularity of the myth, no has ever been successful.

The Cambridge University library is a valuable resource for academics. They have an inventory of over eight million items for researchers' benefit. Each week, the library amasses around 800 entities to add to their catalog. Visitors can access these reserves.
weekly during business hours. Researchers from all fields can find relevant information from the vast supply on the campus.

WC: 59

Adv. Passage 16(C): inventory – catalog— supply – supply

The Cambridge University library is a valuable resource for academics. They have an inventory of over eight million items for researchers' benefit. Each week, the library amasses around 800 entities to add to their catalog. Visitors can access these supply weekly during business hours. Researchers from all fields can find relevant information from the vast supply on the campus.

Adv. Passage 16(D): supply – catalog— inventory – supply

The Cambridge University library is a valuable resource for academics. They have an supply of over eight million items for researchers' benefit. Each week, the library amasses around 800 entities to add to their catalog. Visitors can access these inventory weekly during business hours. Researchers from all fields can find relevant information from the vast supply on the campus.

Adv. Passage 16(A): supply – supply — supply – supply

The Cambridge University library is a valuable resource for academics. They have an supply of over eight million items for researchers' benefit. Each week, the library amasses around 800 entities to add to their supply. Visitors can access these supply weekly during business hours. Researchers from all fields can find relevant information from the vast supply on the campus.
The Cambridge University library is a valuable resource for academics. They have an of over eight million items for researchers' benefit. Each week, the library amasses around 800 entities to add to their. Visitors can access these weekly during business hours. Researchers from all fields can find relevant information from the vast on the campus.