2018-06-01

ESL Students' Reading Behaviors on Multiple-Choice Items at Differing Proficiency Levels: An Eye-Tracking Study

Juan M. Escalante Talavera
Brigham Young University

ESL Students' Reading Behaviors with Multiple-Choice Items at Differing Proficiency Levels: An Eye-Tracking Study

Juan M. Escalante Talavera

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

Master of Arts

Grant Taylor Eckstein, Chair
Troy L. Cox
Steven Luke

Department of Linguistics and English Language

Brigham Young University

ABSTRACT

ESL Students' Reading Behaviors with Multiple-Choice Items
at Differing Proficiency Levels: An Eye-Tracking Study

Juan M. Escalante Talavera
Department of Linguistics and English Language, BYU
Master of Arts

Theorists have been concerned with the overlap of reading and problem-solving for at least a century (Thorndike, 1917, 1973–1974; Sternberg & Frensch, 2014). Various reading models have been proposed, including bottom-up and top-down reading processing (Goodman, 1972; Gough, 1972). In second-language literature, theorists have further noted that reading consists of strategic, purposeful, and interactive processes (Grabe, 2009). In test-taking situations, problem-solving is important because it can compensate for students' language proficiencies. In spite of research showing the use of problem-solving in reading, less is known about how learners actually read and problem-solve in test-taking situations. This study centers around Khalifa, Weir, and colleagues' model for cognitive processing in reading (Weir, Hawkey, Green, & Devi, 2009) in combination with eye-tracking technology in order to examine how ESL readers employ *careful* and *expeditious* reading. Data were gathered from 50 students attending a university-sponsored intensive English program (IEP). Participants read eight validated reading comprehension items at varying difficulty levels while their eye movements were recorded. Results indicate that student level may not be a factor in how carefully and expeditiously a student reads. However, statistical analyses suggest that text difficulty may be a factor in how carefully students read.

Keywords: eye tracking, careful reading, expeditious reading, problem-solving, reading assessment

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to all of those who contributed to this thesis. I am thankful for my chair, Dr. Eckstein, and for the tremendous amount of support, feedback, guidance, and mentorship he provided me as we worked on this thesis together. I am also grateful for the other members of my committee, Dr. Cox and Dr. Luke. Without their support, this project would have been difficult to complete. I am also thankful for my grandparents Mr. and Mrs. James Robert Hladky and for their immense support throughout my educational journey. I would like to thank my parents for sacrificing so much so that my siblings and I could have access to more opportunities in life. Lastly, I am most grateful for my wife, Kathryn. I am thankful for her never-ending support, love, and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This thesis was written in partial fulfillment of the MA TESOL degree at Brigham Young University. As such, the thesis was written in a way that it could be submitted as a manuscript to a journal for publication. *Reading in a Foreign Language* is the journal selected to publish. Those interested in research in reading in a foreign and second language may find this manuscript informative. The length required by the journal is no more than 8,500 words. In preparation for submission we anticipate changes to the current document. Other possible journals include *Language Testing, Journal of Research in Reading,* and *Language Learning Journal.* These journals were selected because they have published related research in the past.

Introduction

Reading in a foreign language has been a well-investigated area over the past several decades (Anderson, 2003; Aslanian-Nemagerdi, 1986; Park, 1997; Park, 2005). Overall, results indicate that high-proficiency readers appear to have a wider repertoire of strategies at their disposal, including being able to use higher-order cognitive skills such as analyzing, reasoning, and elaborating in a second language. On the other hand, less proficient readers tend to rely more heavily on bottom-up strategies and tend to read text and focus on mostly the literal meaning. Another important contribution to the understanding of reading in a foreign language is how common the use of problem-solving strategies appears to be in both proficient and less proficient learners. Theorists have been concerned with the overlap of reading and problem-solving for at least a century (Thorndike 1917, 1973–1974; Sternberg & Frensch, 2014). Various reading models have been proposed including bottom-up and top-down reading processing (Goodman, 1972; Gough, 1972), which ascribe problem-solving to top-down mechanisms. Additionally, Just and Carpenter (1987) as well as Rayner and Pollatsek (1989) developed theories of pre- and post-lexical access and suggested that problem-solving can occur at both locations. In second-language literature, theorists have further noted that reading consists of strategic, purposeful, and interactive processes (Grabe, 2009), that it is a "constellation of interfaced capabilities" including "reasoning and inferencing" (Koda, 2004, p. 227), and that top-down processes are "directed by reader goals and expectations" (Grabe & Stoller, 2002, p. 32). Thus, reading can be conceptualized as a strategic, problem-solving endeavor.

In test-taking situations, problem-solving is important because it can compensate for students' language proficiencies. For instance, Anderson (2003) demonstrated that both proficient and less proficient readers tend to rely heavily on problem-solving when engaged in

online reading. Nevertheless, he also found that more and less proficient readers used the same strategies when taking reading comprehension tests, yet poorer readers used those strategies less successfully (Anderson, 1991). On the other hand, Allan (1992) found that "students are differentially skilled in test taking and that the scores of some learners may be influenced by skills which are not the focus of the test" (p. 101). These skills, collectively called *test-wiseness*, illustrate that problem-solving while reading can affect student performance. Thus, we can see that all learners problem-solve while reading, but they differ in how successfully they do it, and some learners may use problem-solving strategies to compensate for their limitations.

In spite of research showing the use of problem-solving in reading, less is known about how learners actually read and problem-solve in test-taking situations. Current research so far has described test-taking reading processes through retrospective or reflective protocols and item analysis (Bax, 2013; Khalifa & Weir, 2009; Rupp, Ferne, & Choi, 2014; Weir, Hawkey, Green, & Devi, 2009). However, little research has examined readers' test-taking reading behaviors in the moment through eye-tracking measures.

As eye-tracking technology has developed, it has also had a great impact on our understanding of the processes elicited by reading in a foreign language. Eye-tracking has become a popular tool to investigate reading because it essentially allows us to glance into a reader's mind (Rayner, 2009; Rayner, 1998). One of the most recent and well-known eye-tracking studies revealed that there are differences in the use of strategies in a reading assessment between successful and unsuccessful participants (Bax, 2013). Other studies have shown that in a testing environment, readers tend to spend quite a bit of time searching for the correct answer rather than engaging with the text (Wang, Sabatini, O'Reilly, & Feng, 2017). However, research has not investigated differences in problem-solving among readers of different levels. Thus, the

purpose of this research study is to examine how ESL readers employ careful and expeditious reading strategies as they read test items of varying difficulty levels.

Review of Literature

A General Overview of Reading

In order to examine problem-solving, it is expedient to first understand how reading has been defined over the years. Reading comprehension, as described by Bohn-Gettler and Kendeou (2014), involves encoding information, making inferences, and activating background knowledge. Reading has also been defined as a process which is interactive, strategic, purposeful, evaluative, learning, and linguistic (Grabe, 2009). Grellet (1981) explains that understanding a written text is essentially extracting the necessary information as efficiently as possible. McCrudden, Magliano, and Schraw (2010) make the case that not only is reading an intentional act but that it changes from one situation to another. For example, reading in an academic setting involves personal intentions as well as given intentions. In other words, when reading in academic contexts, readers bring personal knowledge and beliefs as well as externally provided instructions about what to do with the text. Furthermore, the case has been made to define reading as "a goal-directed activity in which the person intends to understand a text" (McCrudden & Schraw, 2007). Every student reading in an academic context should set goals to meet the task's demands. These demands could be to summarize an article or to prepare to take the Test of English as a Foreign Language (TOEFL). Reading also involves neurological and cognitive components (Hedgcock & Ferris, 2009). For instance, it has been argued that "reading is a neuronally and intellectually circuitous act, enriched as much by unpredictable indirections of a reader's inferences and thoughts, as by the direct message to the eye from the text" (Wolf, 2007, p. 16). Additionally, it has been proposed that "adept reading is a constellation of interfaced capabilities, ranging from mechanical mappings to more sophisticated conceptual manipulations, such as reasoning and inferencing" (Koda, 2004, p. 227). Bottom-up models have

also been introduced into the reading literature; perhaps one of the most influential models was presented by Gough (1972). In this model, reading is presented as the process by which a reader decodes letters of words, makes grapheme-to-phoneme associations, and constructs meaning from the text. Top-down models have also added to the body of research by proposing that readers are engaged with higher-level processes and that hypothesis testing and checking is a characteristic of fluent readers (Goodman, 1972). Weir and Khalifa (2008, 2009) conceptualized reading as a multicomponential construct. In their model, reading can be broken down into strategies used to accomplish the given task. They added that reading construct needs to be described by observing the array of strategies readers use to construct meaning and complete the tasks. Based on these definitions of reading, it is evident that reading is a complex construct and that such complexity may have an impact on how ESL students read.

Reading as Problem-Solving

Over the last century, various reading models have been proposed as an attempt to distinguish between reading and problem-solving (Thorndike 1917, 1973–1974; Sternberg & Frensch, 2014; Goodman, 1972; Gough, 1972). Even though the bottom-up model presented by Gough (1972) has been labeled as being in "complete contrast to the top-down frameworks" (Sternberg & Frensch, 2014, p. 6), these two reading models agree that reading shares similarities with problem-solving. For instance, it has been argued that Gough's (1972) model presents processes that could be performed without requiring much cognitive capacity (Sternberg & Frensch, 2014). Additionally, in regard to Goodman's (1972) top-down model, although it elicits higher-level processes, it too resembles problem-solving (Sternberg & Frensch, 2014).

In a more recent model, Khalifa and Weir (2009) made the distinction between three main components: metacognitive activity, the central processing core, and the knowledge base

(Brunfaut & McCray, 2015). Metacognitive activity involves the reader deciding the types of reading necessary to complete a given task. These types of reading include *local, global, careful,* and *expeditious.* The central processing core is composed of eight cognitive processes that work together and are thought to lead to reading comprehension. These cognitive processes are divided into lower level and higher level. The lower level includes word recognition, lexical access, syntactic parsing, and establishing propositional meaning. On the other hand, higher-level processes include inferencing, building a mental model, creating at the text level, and intertextual representation (Brunfaut & McCray, 2015; Khalifa & Weir 2009). Finally, the knowledge base component is essentially sources of knowledge the reader may possess; these may include knowledge of word orthography, phonology, and morphology, as well as lexical knowledge, syntactic knowledge, and knowledge about the subject, to name a few (Brunfaut & McCray, 2015). Due to the narrow scope of this study, we focus only on the metacognitive activity component. According to the reading model, readers set goals as they begin to interact with a text. This means they must decide which types of reading are more appropriate in order to accomplish the tasks at hand. These types of reading include *local* versus *global* and *careful* versus *expeditious* reading. *Local* reading takes place at the sentence level. It involves word recognition, lexical access, and synthetic parsing and establishing explicit propositional meaning at the phrase, clause, and sentence level. *Global* reading, on the other hand, deals with the comprehension of main ideas in a text. It entails going beyond the sentence level and understanding the connections between ideas in a text (Bax & Weir, 2012).

According to Khalifa and Weir's (2009) model, *careful* reading entails extracting complete meaning from text either at the local or global level. This type of reading is based on reading for comprehension, which is slow, careful, linear, and incremental. *Expeditious* reading,

in contrast, involves quick, selective, and efficient reading in order to gain access to desired information in a text. This type of reading likely includes *skimming, scanning,* and *search reading.* Similarly to *careful* reading, *expeditious* reading can be present at the local or global level (Bax & Weir, 2012; Brunfaut & McCray, 2015; Khalifa & Weir, 2009). This model of reading has been used extensively to investigate various issues associated with second-language testing, including construct validity of the International English Language Testing System (IELTS), cognitive processes employed by test takers, and the validity of multiple-choice reading tests (Bax & Weir, 2012; Brunfaut & McCray, 2015; Katalayi, & Sivasubramaniam, 2013; Weir, Hawkey, Green, & Devi, 2009). In this study we will examine careful and expeditious reading at the global level.

Measuring Careful and Expeditious Reading

We have chosen to investigate careful and expeditious reading through eye-tracking methodologies and measures. These include first- and second-pass dwell time durations and look-back count measures. First-pass dwell time refers to the total amount of time in milliseconds a reader spends on a specific area of interest (AOI) before moving on to a second AOI (Holmqvist, 2011). This measurement has been linked to early processing (Liversedge, Paterson, & Pickering, 1998; Loftus & Mackworth, 1978). Second-pass dwell time refers to the sum of all returning dwell times to an AOI (Hyönä, Lorch, & Rinck, 2003). Look-backs, on the other hand, identify the number of times a reader returns to a specific AOI (Holmqvist, 2011). Clifton, Staub, and Rayner (2007) elaborate on look-backs by stating that this measurement can indicate effort by the reader to coordinate information and solve a problem. This rereading behavior is indicative of *expeditious* reading; in other words, readers are not necessarily reading for comprehension but rather to answer a multiple-choice question, for example. Nation (2009),

argued that in normal skilled reading, 300 words per minute (wpm) tends to be an indication of *careful* reading. He also elaborated that if an individual is reading at a rate of 400 wpm, he or she can no longer be considered to be doing a *careful* reading but rather *expeditious* reading. Research has shown that speed-reading can transfer to other types of reading and that it does not affect comprehension (Macalister, 2010; Tran & Nation, 2014). Whereas reading is frequently measured in wpm, this study takes a different approach. Taking into consideration Khalifa and Weir's (2009) reading model, Nation's (2009) measurements of *careful* and *expeditious* reading, as well as Holmqvist's (2011) and Conklin, Pellicer-Sánchez, and Carrol's (2018) definitions of first- and second-pass dwell time and look-backs, we propose to use total dwell time as a proxy for *careful* reading and look-backs as a measurement of *expeditious* reading.

Reading at Different Proficiency Levels

Reading for learners of English, as Anderson (2003) explained, is the most important skill to master in order to ensure future success in language learning. He further elaborated that when learners of English have strong reading skills, they tend to make great progress in other areas of the target language. In 2003, Anderson carried out a study in which he investigated the online reading strategies of second-language learners in English-as-a-foreign-language (EFL) and English-as-a-second-language (ESL) contexts, ranging in proficiency from high beginning to high intermediate. He came to the conclusion that one of the most used strategies by language learners is problem-solving.

Other studies have investigated the relationship between learners' proficiency level and reading strategies. In 2005, Park investigated whether second-language learners from various proficiency levels varied in the reading strategies they employed. By means of recall protocols, Park found that when reading the same text, high-proficiency ESL students read better than low-

proficiency ESL students. In her study, Park challenged the notion that low-proficiency level ESL students use bottom-up strategies more frequently than top-down strategies. She stated that "lower proficiency students also used more top-down strategies than bottom-up strategies when they read in an L2" (Park, 2005).

Aslanian-Nemagerdi (1986) investigated how proficient and less-proficient ESL students dealt with reading and processing texts. In this study, it was concluded that skilled ESL readers often integrated background knowledge with the text, made inferences while reading, and successfully overcame text difficulties by reading on or rereading. On the other hand, less proficient readers tended to stay close to the literal meaning of the text, rarely activated background knowledge, and used dictionaries or translators. Aslanian-Nemagerdi closed the study by stating that in order to become proficient readers, ESL students must acquire text schemata as well as knowledge schemata. In other words, it seems that more proficient readers could do both, that is, read carefully and expeditiously, whereas less-proficient readers were mainly only able to read carefully.

Park (1997) investigated the relationship between learning strategies and second-language proficiency among Korean university students. The subjects' proficiency level ranged from intermediate to advanced. Learning strategies were measured by the Strategy Inventory for Language Learning (SILL) and L2 proficiency was determined by the TOEFL. The study also examined whether such a relationship is linear or curvilinear. The findings indicated that the relationship between learning strategies and proficiency is linear. This means, according to Park (1997), that as students' proficiency level increases, so does the range of learning strategies implemented. According to this study, the most predictive learning strategies of the TOEFL scores were cognitive and social strategies. As we interpret these results, we can come to the

conclusion that active mental engagement with the text is needed to engage in higher order cognitive processes such as analyzing, reasoning, and elaboration (Park, 1997). As we can see, reading takes different forms and its construct is reshaped depending on the reader's skills and proficiency level. Reading in a test-taking environment has also been extensively investigated.

Cognitive Processing in Test-Taking Environments

Traditionally, researchers aiming to glance into readers' cognitive processes have relied mainly on retrospective protocols (Bax & Weir, 2012; Brunfaut & McCray, 2015; Weir, Hawkey, Green, & Devi, 2009). With improvements to eye-tracking technology, a new door has opened up. Recent studies have used both eye-tracking technology and retrospective protocols to gain insights into readers' actual cognitive processing as opposed to what they report (Bax & Weir, 2012).

There seems to be a wide array of possibilities when incorporating retrospective protocols and eye-tracking into a study's design. For example, there have been instances where the retrospective questionnaire appears after each test item is completed. This timing is meant to elicit from participants their interpretation of how they engaged with the test item (Bax & Weir, 2012). Other studies have developed questionnaires meant to gain insights into cognitive processes in academic reading. These surveys are administered upon completion of testing tasks (Weir et al., 2009). Furthermore, other scholars have adopted what is known as *stimulated recall methodology.* This methodology uses eye traces as stimuli to elicit participants' cognitive processes (Brunfaut & McCray, 2015).

Results from these retrospective protocols appear to validate their use as complementation to eye-tracking technology. For instance, Bax and Weir (2012) found that participants had been accurate in their self-reports 68.4% of time. These results may provide

evidence to claim that the design of the study was effective (responding after every test item rather than waiting until the end of the test). Weir et al. (2009) found that participants completing IELTS test items appear to have an extensive scope for careful reading. This finding is due to the way the test is structured. The test provides multiple opportunities to read the text several times before finding the necessary information to respond. Brunfaut and McCray (2015) reported similar results. They arrived at the conclusion that in a testing environment, readers mostly adopt *careful* reading over *expeditious* reading. They also conclude by stating that this methodology could be valuable to test validation research.

Eye-Tracking and Test-Taking While Reading

English learners' reading ability is assessed extensively in the form of high stakes test such as the TOEFL and other language proficiency exams. In this section, we present various studies that have examined language learners' behaviors while reading in a test-taking environment.

Eye-tracking technology has been used extensively to investigate reading behaviors in English-language learners. Rayner, a cognitive psychologist best known for being a pioneer in eye-tracking research in reading, made the case for eye-tracking as a tool for better understanding reading (Rayner, 1998; Rayner, 2009). He presented an exhaustive review of previous studies that have used eye-tracking to investigate various reading tasks (Rayner, 1998). In this particular review of studies, Rayner advocated for eye-tracking due to the fact that it has proven to be informative in revealing moment-to-moment processing while reading. Almost a decade later, he reaffirmed his position on eye-tracking research by stating that eye movements give us a very precise indicator of mental processing (Rayner, 2009). In general, eye-tracking technology has enabled researchers to investigate reading. Some of the studies presented in this

section range from the influences of font size and type on online reading (Beymer, Orton, & Russell, 2007) to a comparison of silent and oral reading (Ashby, Yang, Evans, & Rayner, 2012) to morphological processing in reading (Bertram, 2011).

Wang, Sabatini, O'Reilly, and Feng (2017) showed that it is not only individual differences that can affect reading behaviors but also task requirements. In their study, which investigated precisely that, they found that task requirements also have an impact on reading behaviors. They found that when readers were given two distinct tasks after reading—to answer a multiple-choice question or write a summary—they spent significantly less time reading when they knew they would have to answer to a multiple-choice question. The conclusion here is that in a testing environment, students appear to spend time "searching for answers" rather than engaging in coherence. Other studies have shown that reading is affected not only by while-reading and post-reading strategies but also by pre-reading strategies (Guo, Kim, Yang, & Liu, 2016). Their study sought to investigate exactly how effective previewing answer options in a multiple-choice test can be. Their results showed that "negative effect of previewing on response accuracy and response time was found for higher-proficiency participants but not for low intermediate learners" (Guo et al., 2016). The interpretation of these results is that high-proficiency readers have the mental capacity to engage in higher-order thinking processes.

Bax (2013) focused on investigating test takers' cognitive processing while completing an onscreen IELTS reading test. The study's focus was to analyze the differences between successful and unsuccessful test takers. After recording eye-movements of a random sample of participants (n= 38), Bax found that although only five of the ten test items analyzed showed significant differences, it was possible to differentiate between successful and unsuccessful test takers. According to the study, the two groups of test takers differed in their cognitive processing

at the lexical and grammatical levels as well as regarding expeditious reading and metacognitive awareness. Bax suggested that it is possible for test designers to target several specific cognitive processes based on the results from the study and that some test items were able to distinguish successful and unsuccessful test takers. He stated that there are several implications for teachers and students. First, because successful test takers showed the ability to read expeditiously, students and teachers should work on developing reading strategies that will allow them to find the location of correct answers as quickly as possible. Second, because successful participants showed greater abilities in lexical knowledge, it is appropriate to suggest that students should work on developing lexical knowledge. Finally, the study provided evidence to suggest the need for learning to deal with grammatical ambiguity in tests because ambiguity can make finding the correct answer challenging.

Wang et al. (2017) investigated how readers from different reading proficiency levels interacted differently with reading tasks. They recruited 60 native English-speaking undergraduate students from a public university in the United States. In order to measure the participants' reading efficiency, they used the Efficiency of Basic Reading Comprehension (EBRC) task (Sabatini, Bruce, Steinberg, & Weeks, 2015). Participants then read four passages, and two conditions were presented. In the first condition, readers were given the task of writing a summary before answering multiple-choice questions. In the second condition, readers were told to work directly on multiple-choice questions without writing the summary. This study used the Tobii T60 eye-tracking system with a sampling rate of 60 Hz. The researchers also identified three AOIs: the reading passages, the multiple-choice questions (including the stem and options), and a summary writing area. The results showed that, overall, in a testing environment, students spend more time "searching for answers" rather than engaging deeply with the text. Results also

showed that high-proficiency readers appear to have more strategies to process the texts faster. Furthermore, the study showed that even low-proficiency readers benefitted from having to write a summary prior to answering multiple-choice questions. Unfortunately, these promising results did not transfer over when the summary was not required.

Text Difficulty

Other studies have provided evidence that text difficulty has virtually no negative impact on L2 reading comprehension (Chiang, 2016; Bahmani & Mohammad, 2017) when readers are involved in extensive reading (ER). Chiang (2016) focused on examining the relationship of varying text difficulty on L2 reading attitudes and reading comprehension. Regarding the effects on reading comprehension, she found that all participants significantly improved regardless of the text difficulty. Furthermore, Bahmani, and Mohammad (2017) investigated the effects of text difficulty on reading anxiety and reading comprehension. Their results show that, regardless of text difficulty, participants improved their reading comprehension. Additionally, research has shown that in terms of reading multiple-choice question stems, readers appear to use them to dictate how expeditiously they read (Harasym, Price, Brant, Violato, & Lorscheider,1992; Yanagawa & Green, 2008). These studies were similar in that both used the Oxford Bookworms Library Series published by Oxford University Press as their reading materials. Both studies used similar methodology by randomly diving participants into two groups (i.e., *i+1* and *i-1*). In Chiang's (2016) study, all of the participants were high-intermediate, while Bahmani and Mohammad's (2017) participants scored into level A2—low according to the Common European Framework of Reference (CEFR).

Despite research on reading and, more specifically, on testing environments (Bax, 2013; Bahmani, & Mohammad, 2017; Chiang, 2016; Wang et al., 2017), questions still remain

unanswered as far as exactly how readers of varying levels interact with reading passages and stems of varying difficulty. Bax (2013) and Wang et al. (2017) provided valuable contributions to the understanding of reading behaviors while taking tests. However, their studies lacked information on how the students' level affected their reading behaviors when interacting with the stem and text of varying difficulties.

Research Questions

In this study, we are interested in investigating what happens to first- and second-pass dwell times and look-backs as the readers' level increases and the difficulty level of test items also increases. Based on this interest, we have created the following research questions:

1. How do students at different levels expeditiously read texts of different difficulty levels in terms of passage look-backs and stem look-backs?
2. How do students at different levels carefully read texts of different difficulty levels in terms of passage dwell times and stem dwell times?

Research Design

Participants

The participants for this study were 50 English-language learners from an intensive English program (IEP) associated with a large university in the western United States. Participants were recruited from pre-existing courses: 14 came from Level 2, 17 from Level 4, and the final group of 19 came from Level 6, with Level 6 being the highest level at this school. This IEP associates class levels with the American Council of the Teaching of Foreign Languages (ACTFL) proficiency guidelines. For example, Level 2 roughly corresponds to intermediate-low in the ACTFL scale; Level 4 corresponds to intermediate-mid; and Level 6 corresponds to advanced-low. However, a student in Level 2, for example, may range in proficiency in terms of speaking, listening, reading, and writing. Furthermore, class placement is decided by taking into consideration all the aforementioned language skills. This placement strategy means that in a single class level, reading proficiency can be quite diverse. Thus, it is not appropriate to assume all participants from Level 2 possess the same reading proficiency. The 50 participants were evenly distributed by gender: 25 males and 25 females. The age of the participants ranged from 18 to 47 years old, with the average of all three groups being 26.7 years old. In order to control for language, all the participants were native Spanish speakers learning English as their second or additional language, and they came from the following ten countries: El Salvador, Honduras, Mexico, Argentina, Bolivia, Chile, Ecuador, Peru, Venezuela, and Spain. In addition, the participants' English-language learning backgrounds varied. For example, the length of time these students had been enrolled in their current intensive English language program ranged from 2 to 14 months in length. Similarly, the time they had spent studying

English in their native countries varied, ranging from no instruction at all to as much as 13 years (1 hour per week, approximately).

Materials

The materials used in this experiment were eight reading comprehension test items. The reading passages were developed by the Center of Language Studies at Brigham Young University and had been previously administered with the item difficulty parameters aligning to expected proficiency levels, thus providing validity evidence for their use in a proficiency test. The selected reading passages represented three levels of ACTFL proficiency guidelines: intermediate, advanced, and superior. The topics of the reading passages at the intermediate level included newspaper ads, telephone notes, and signs for tourists, and the average word count was 31. At the advanced level, passages contained news articles that communicated straightforward, factual information as shown in Figure 1. The superior-level passages contained political commentary, analysis, and opinion editorials. The word count average between these two levels was quite similar, 155 for advanced and 158 for superior. Each passage was accompanied by a single multiple-choice question that aligned reader task with the author's purpose. The question had five possible response options with one clearly correct answer, three distractors, and the fifth option labelled as "I don't know."

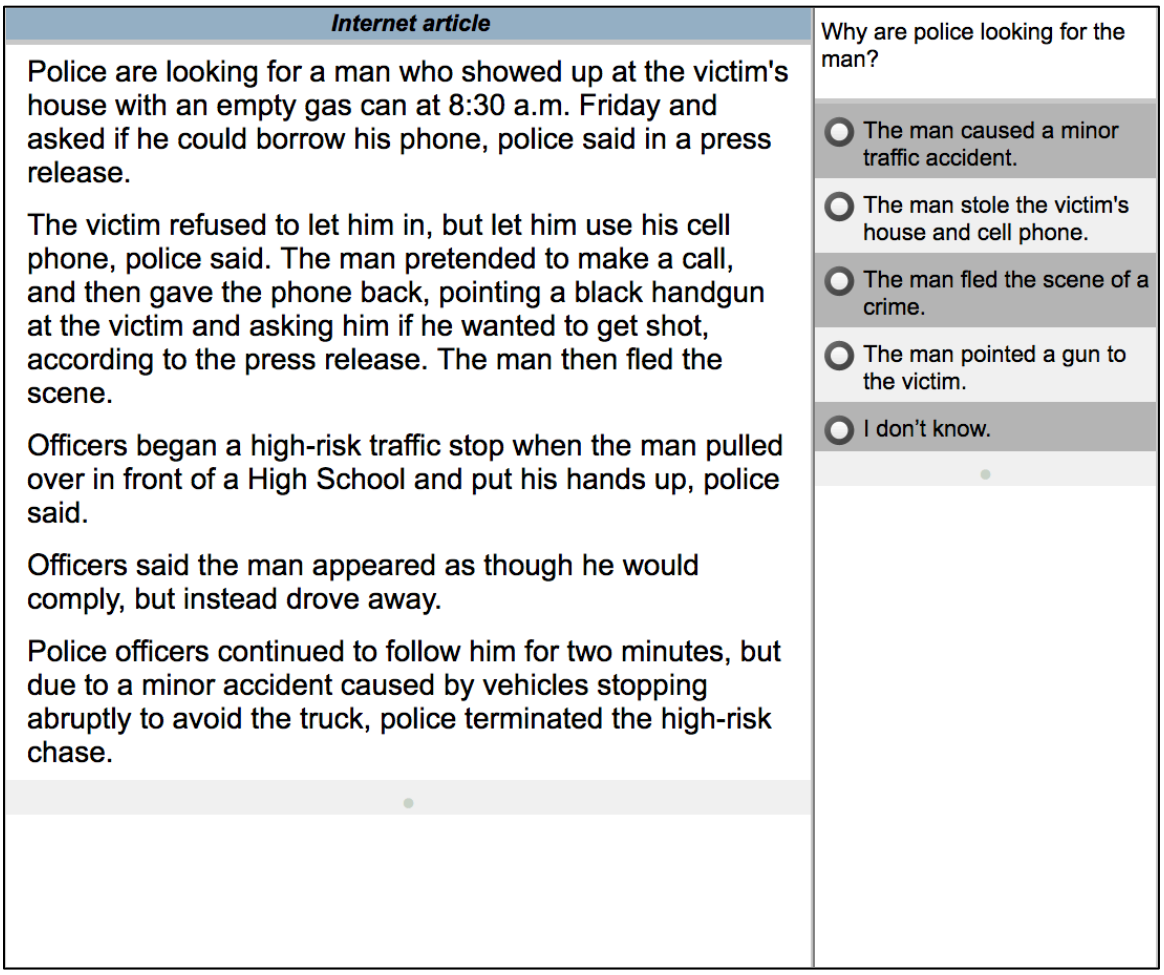


Figure 1. Example of advanced-level reading passage used in this study.

Apparatus

To measure eye movements for this study, an SR Research Eyelink 1000 Plus eye-tracker with a 35 mm lens and a sampling rate of 1000 hz (1000 measurements per second) was used. As participants performed the tasks, they used a chin rest to reduce head movement and they sat 65 cm away from a 21-inch computer monitor where the reading passages were being displayed. Participants also used a computer mouse to select their responses to the multiple-choice questions. It is important to mention that prior to (and during, as necessary) the experiment, 9-point calibrations were performed with every participant to ensure the collection of accurate data.

Procedures

To ensure uniformity in format, screenshots of the reading passages and their accompanying question were programmed into our eye-tracker. This uniformity in format was crucial. As we used screenshots, we controlled for font size, location, dimensions, and we were able to match an online testing situation. Uniformity was also needed to limit programming difficulties. We divided each passage into seven different AOIs, one around the reading passage, one around the stem, and one for each of the five response options. These AOIs were essential to the collection of data because the eye-tracker records subjects' eye movement behavior for each preselected AOI, thus making it possible to interpret reading behaviors. We decided not to code each word as an AOI for this study. Although word-level AOIs would be a useful method for collecting data in a follow-up study, the aim of the present study was to analyze reading behaviors at the global level.

Prior to the experiment, participants gave informed consent to be research subjects as per institutional review board (IRB) requirements. They were then positioned in front of the eye-tracking machine where each participant's head was stabilized and a 9-point calibration was performed. Participants were given verbal instructions prior to beginning the reading task. They were told they would read eight short reading passages (one was removed from the study) and that each passage would contain a multiple-choice question. They were instructed to read as naturally as possible. They were also encouraged to reread if necessary. In this study, we wanted to collect data that resembled natural test reading as closely as possible. The reading passages and accompanying question were presented one at a time and in an increasing order of difficulty, intermediate level first, followed by advanced and superior. After each reading, subjects were instructed to fixate on a dot in the middle of the screen for drift correction. If the eye-tracker was

not able to accurately track the subject's eye, a recalibration took place. After the participants had read five of the nine passages, they were given the option to take a short break. If they opted to take the break, the eye-tracker was recalibrated prior to resuming the experiment. In average, the experiment took approximately 30 minutes.

Data Measurements

Although a great amount of data were gathered by the eye-tracker, this study focuses on examining the readers' interaction with the text and stem in terms of first-pass dwell time, second-pass dwell time, and look-backs. First-pass dwell time is the term used to refer to the total amount of time (in milliseconds) an individual dwelled on an AOI prior to moving on to a second AOI (Holmqvist, 2011). First-pass run time has been associated with early processing and object recognition (Liversedge, Paterson, & Pickering, 1998; Loftus & Mackworth, 1978). Second-pass dwell time has been defined as the sum of all returning dwell times to an AOI (Hyönä, Lorch, & Rinck, 2003). Look-backs are used to identify participants who regress into a specific AOI (Holmqvist, 2011). Figure 2 further elaborates on these definitions. In this figure, a hypothetical reading pattern is portrayed, and it can be noted that the first-pass dwell time in AOI 1 consists of two fixations. Likewise, the first-pass dwell time on AOI 2 only has two fixations. The second-pass dwell times for AOI 1 and 2 are also identical with just a single fixation. As the figure illustrates, a regression occurs when a participant returns to a particular AOI and dwells there for a period of time.

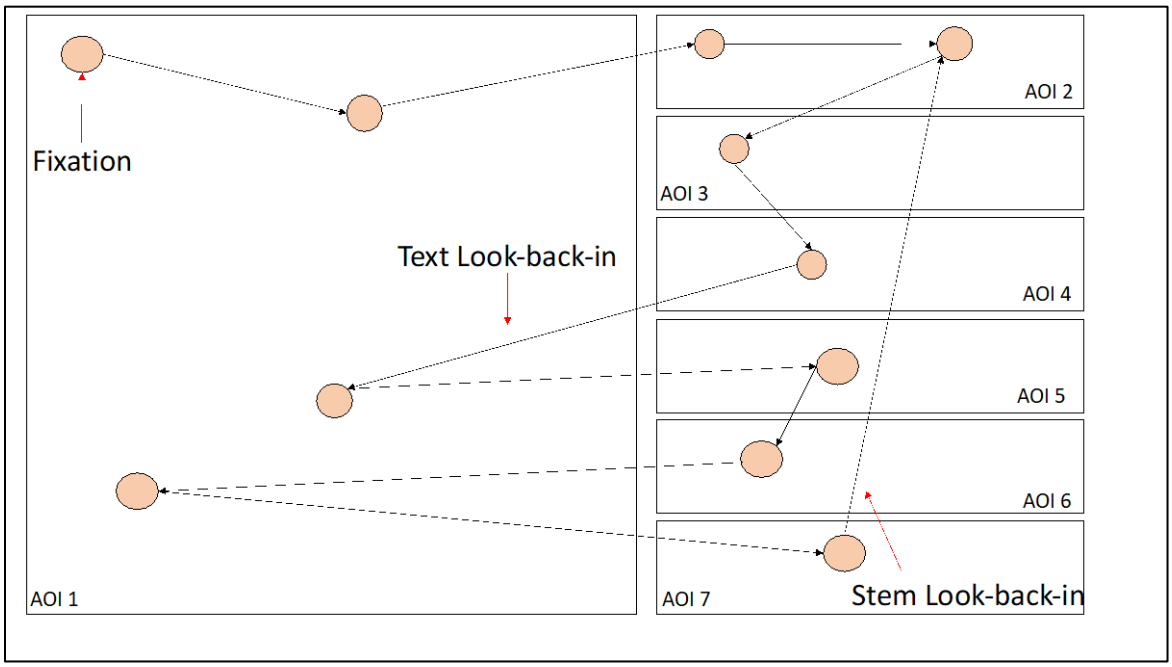


Figure 2. Scan path between AOIs.

This study adopts Khalifa and Weir's (2009) and Weir et al.'s (2009) model for cognitive processing in reading. This model conceptualizes reading into four dimensions: *expeditious* versus *careful* reading and *local* versus *global* reading. Because of time and requirement constraints, this study focuses on *expeditious* and *careful* reading at the global level only. *Expeditious* reading is defined as "quick, selective and efficient reading to access relevant information in a text." *Careful* reading, on the other hand, is "slow, careful, linear, incremental reading for comprehension" (Khalifa & Weir, 2009; Weir et al., 2009). We have paired these dimensions with eye-tracking measures based on scholars' conceptualizations of them (Holmqvist, 2011; Paterson & Pickering, 2003; Loftus & Mackworth, 1978).

We decided to use first- and second-pass dwell time as two of our variables only after carefully considering research done up to this date. First, based on the definition provided by Holmqvist (2011); Liversedge, Paterson, and Pickering (2003); and Loftus and Mackworth (1978), it appeared as an obvious choice to include first- and second-pass dwell time data. We

wanted to know how long participants spent dwelling on text passages and stem questions of varying difficulties. We incorporated Khalifa and Weir's (2009) reading model, which presents *careful* and *expeditious* reading. We decided to use first- and second-pass dwell time as measurements of *careful* reading. Additionally, look-backs are indicative of rereading behavior between AOIs. This rereading behavior appears to complement well what Khalifa and Weir (2009) referred to as *expeditious* reading. Thus, in this study, we use first- and second-pass dwell time to measure *careful* reading and look-backs to measure *expeditious* reading.

Using these measures for reading, this study aims to add to the understanding of foreign-language reading in a testing environment. This study incorporated a component that previous studies had not investigated thus far. We had three groups of readers of varying levels interacting with reading assessment items of three differing difficulty levels. The results of this study aim to add to the extensive body of research on reading.

Data Analysis

Multiple within-measures analyses of variance (ANOVAs) were used in order to address the study's research questions. ANOVAs are commonly used to compare means between two or more independent variables. Our study was composed of three dependent variables (first-pass dwell time, second-pass dwell time, and look-backs) and two independent variables (test item difficulty and student proficiency). The ANOVA allows each dependent variable to be compared with each independent variable. With this analysis we were able to see which independent variables had an effect on our dependent variables. We ran the analysis on IBM SPSS Statistics 25.

Results

The focus of this study was to examine how English learners of various language proficiencies engage with reading tests of varying difficulties. The research aims were as follows: (1) How do students at different levels expeditiously read texts of different difficulty levels in terms of passage look-backs and stem look-backs? (2) How do students at different levels carefully read texts of different difficulty levels in terms of passage dwell times and stem dwell times?

Specifically, we wanted to see if students showed differences in their dwell times on passages of texts with differing difficulty levels. Likewise, we wanted to see if their look-backs behaved similarly. We also wanted to see if these behaviors changed for the stem of each comprehension question.

Expeditious Reading

Passage look-backs.

We examined the number of look-backs students made while reading the entire reading passages, including all look-backs students made during each trial. The analysis of the effects of student level and text level on text look-backs generated the following results. Student level $F(2, 48) = .608$, $p=.548$ and text level $F(2, 96) = 1.271$, $p=.282$ were not statistically significant. The interaction effect between student level and text level was significant, $F(4, 96) = 6.891$, $p <.001$. The collected data for student proficiency by text difficulty are shown in Table 1.

Table 1. *Average Text Look-Backs for Student and Difficulty Levels*

		Text Difficulty						
		N	Intermediate	SD	Advanced	SD	Superior	SD
Student Level	Level 2	14	3.91	2.028	2.97	1.855	3.35	2.755
	Level 4	17	2.52	1.333	3.41	1.706	2.57	1.403
	Level 6	19	2.21	1.262	2.78	1.483	3.68	2.610
	Total	50	2.79	1.663	3.05	1.657	3.20	2.322

We noticed some interaction patterns from these data set. First, fewer look-backs were made with the intermediate-level reading passages at 2.79. However, the Level 2 group made 3.91 look-backs when reading text at this difficulty level. Additionally, we noticed that the look-backs at the advanced difficulty level appear to be closer to one another than at the other difficulty levels. In terms of student proficiency, the Level 6 group exhibited a more linear pattern in their look-backs. As difficulty increased, more look-backs were made. These data are displayed in Figure 3.

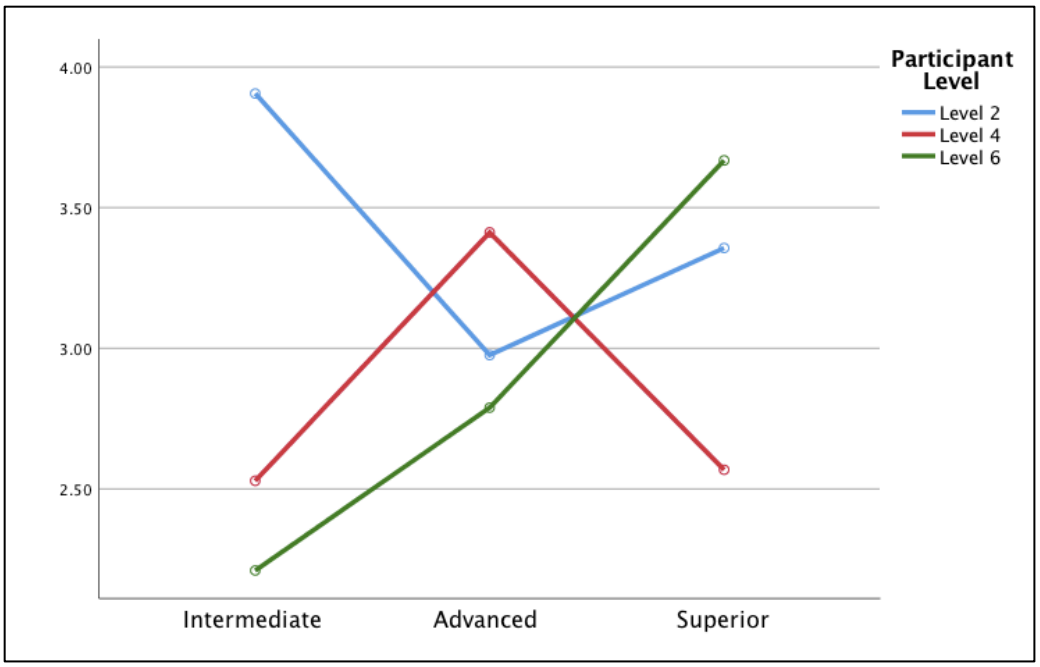


Figure 3. Text regression changes across passage level by student level.

Stem look-backs.

We also analyzed the number of look-backs students made on the question stems. This analysis included the total number of look-backs made on all of the reading tasks. The ANOVA calculated on stem look-backs revealed that student level $F(2, 47) = 1.018$, p=.369, text level $F(2, 94) = 2.038$, p=1.47 were not significant. The interaction effect between student level and text difficulty was also not statistically significant, $F(4, 94) = .848$, p= .476. The collected results for student proficiency by text difficulty are displayed in Table 2.

Table 2. *Average Stem Look-Backs for Student and Difficulty Levels*

			Text Difficulty					
		N	Intermediate	SD	Advanced	SD	Superior	SD
Student Level	Level 2	14	1.36	.529	1.05	1.037	1.32	.745
	Level 4	17	1.02	.651	0.79	.587	1.26	.507
	Level 6	19	1.29	.909	1.61	2.216	1.91	2.855
	Total	50	1.22	.733	1.18	1.527	1.52	1.822

A closer look at the data revealed a few interaction patterns. For example, the most total look-backs were made while reading the superior-level question stems. As we look at the total number of regression counts made on each difficulty level, we can see that that difference is quite narrow. In terms of student-level performance, we can see that once again, the Level 6 group read the stems in a linear pattern; that is, as stem difficulty increased, so did the number of look-backs. Our data also shows that the Level 2 and Level 4 groups exhibited very similar patterns. The only observable difference is that the Level 2 group made more look-backs on all the stem difficulty levels. Figure 4 displays these data.

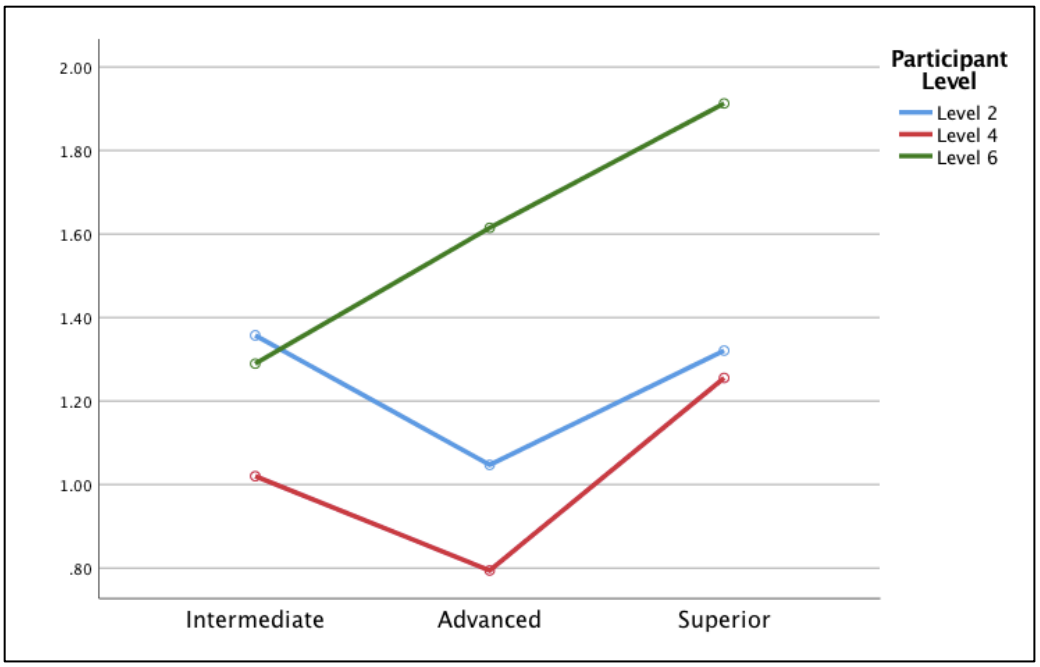


Figure 4. Stem look-back changes across passage level by student level.

Careful Reading

Passage first-pass dwell time.

We examined the amount of time in milliseconds participants read the entire text structure the first time they encountered it. This examination included all fixations across the first run as students interacted with the reading text only. An ANOVA was calculated on the text first-pass dwell time. Although descriptive data showed differences, the analysis shows that student level $F(2, 47) = .917$, $p = .407$ was not significant but the text level $F(2, 94) = 5.666$, $p = .007$ was statistically significant. The interaction effect between student level and text difficulty was not statistically significant, $F(4, 94) = .521$, $p = .699$. The collected results for student level by text difficulty are displayed in Table 3.

Table 3. *Average Text First-Pass Dwell Time for Student and Difficulty Levels in Milliseconds*

		Text Difficulty						
		N	Intermediate	SD	Advanced	SD	Superior	SD
Student Level	Level 2	14	8,610.14	5,071.99	16,021.07	11,874.96	13,914.07	11,953.73
	Level 4	17	7,809.00	5,550.00	13,741.88	19,747.96	18,116.82	24,981.59
	Level 6	19	5,272.42	3,492.11	11,558.37	13,103.87	9,819.26	14,433.42
	Total	50	7,069.42	4,845.91	13,550.32	15,202.77	13,786.98	18,188.77

Two important patterns emerged from the descriptive results. As might be expected, students tended to have higher reading times as the text increased in difficulty. For instance, among the intermediate-level texts, students on average read for just over 7 seconds, whereas with the superior-level text, readers read nearly twice as long at about 13.7 seconds. We also noticed a pattern with the intermediate- and advanced-level passages and student level. As the student level increased, less time was spent dwelling on the passages. These numbers are displayed in Figure 5. They show that Level 2 and Level 6 had very similar reading patterns. Both groups tended to read the text for slightly more time as these readings progressed in difficulty, and once they read at the superior level, their first-pass dwell time decreased by about 3 seconds. The Level 4 students, on the other hand, read in a linear pattern. As passage difficulty increased, their first-pass dwell time also increased.

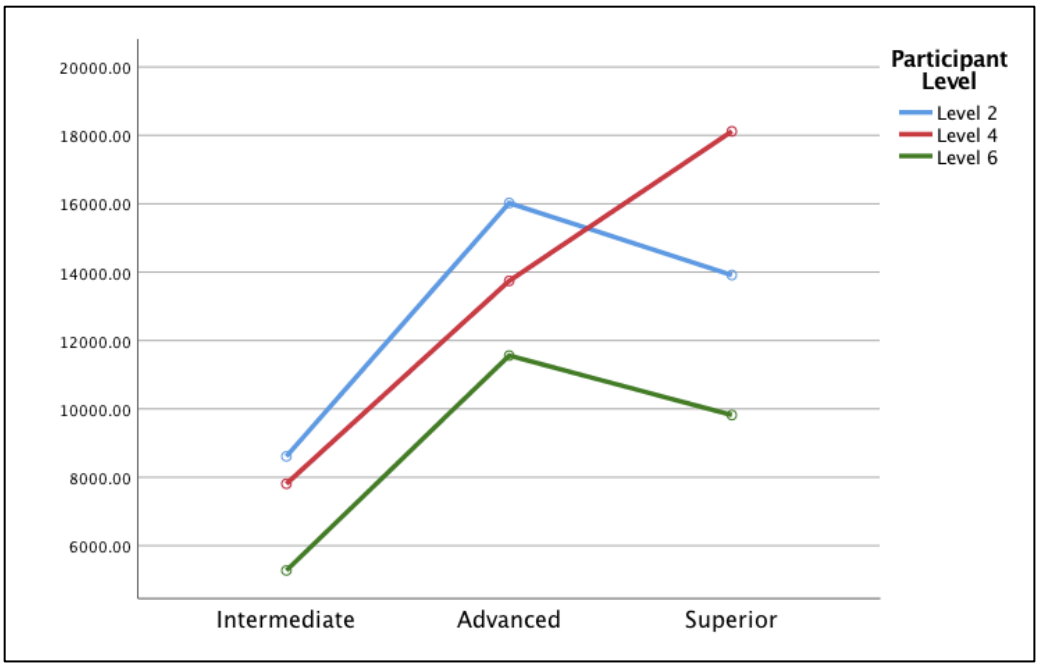


Figure 5. Text first-pass dwell time changes across passage level by student level.

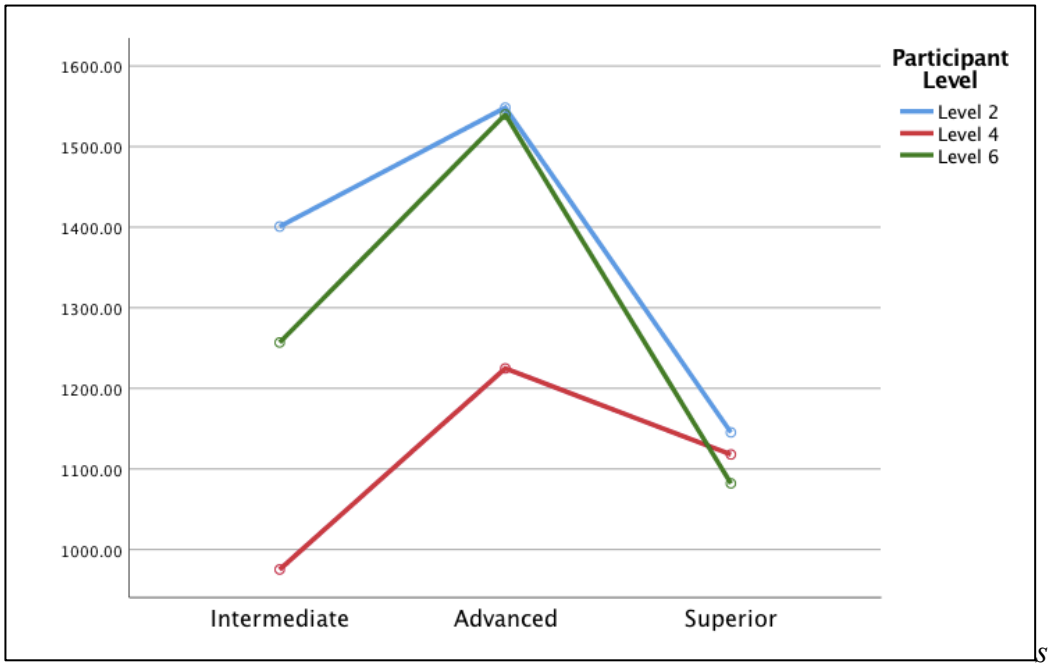
Stem first-pass dwell time.

We also examined the amount of time students spent reading the question stems. This was composed of fixations from the first run as students read the stems only. The ANOVA performed to analyze stem first-pass dwell time revealed that student level $F(2, 47) = .98$, $p = .410$ was not significant but the text level $F(2, 94) = 3.481$, $p = .047$ was statistically significant. The interaction effect between student level and text difficulty was not statistically significant, $F(4, 94) = .607$, $p = .620$. Table 4 presents the results for student level by text difficulty.

Table 4. *Average Stem First-Pass Dwell Time for Student and Difficulty Levels*

			Text Difficulty					
		N	Intermediate	SD	Advanced	SD	Superior	SD
Student Level	Level 2	14	1,400.79	581.358	1,548.29	1,513.391	1,145.29	743.780
	Level 4	17	975.24	630.027	1,224.76	791.267	1,117.94	488.054
	Level 6	19	1,256.89	435.078	1,539.74	769.683	1,082.05	540.693
	Total	50	1,201.42	564.901	1,435.04	1,026.139	1,111.96	576.731

A couple of patterns can be noted from these results. First, in terms of text difficulty, the intermediate-level question stems were read the fastest at just over 1 second. On the other hand, the advanced-level question stems were read a bit longer at close to 1.4 seconds. It is also worth mentioning that the intermediate-level and superior-level stems were read at nearly the same rate. In terms of student proficiency, the Level 2 and Level 6 groups behaved in very similar patterns, spending more time reading the advanced-level stems and showing a sharp decrease as the superior-level stems were read. These numbers are displayed in Figure 6.



s

Figure 6. Stem first-pass dwell time changes across passage level by student level.

Passage second-pass dwell time.

We examined the amount of time in milliseconds participants read the entire text structure the second time. This analysis included all fixations across the second run as students interacted with the passage only. An ANOVA was calculated on the text second-pass dwell time. The analysis shows that student level $F(2, 47) = 1.259$, $p = .293$ was not significant and the text level $F(2, 94) = 14.159$, $p < .001$ was statistically significant. The interaction effect between student level and text difficulty was not statistically significant, $F(4, 94) = 1.236$, $p = .301$. The collected results for student level by text difficulty are displayed in Table 5.

Table 5. *Average Text Second-Pass Dwell Time for Student and Difficulty Levels in Milliseconds*

			Text Difficulty					
		N	Intermediate	SD	Advanced	SD	Superior	SD
Student Level	Level 2	14	5,914.29	4,021.698	22,793.93	12,079.970	13,651.86	13,883.168
	Level 4	17	4,995.18	4,513.176	17,297.82	26,798.886	21,691.12	16,481.684
	Level 6	19	3,758.58	3,758.578	13,170.89	9,998.387	15,556.74	12,825.598
	Total	50	4,782.62	4,782.620	17,268.50	18,032.733	17,109.06	14,556.679

A few patterns emerged from the descriptive results. First, readers appear to have spent less time reading the intermediate-level passages the second time around. Our data shows that they dwelled on the passage for approximately 4.7 seconds, whereas the first time they dwelled for around 7 seconds. The data also suggests that as student level increased, less time was taken to read the intermediate and advanced-level passages. This is only true for Level 2 readers while reading the superior-level passages. In addition, our data shows that Level 4 and Level 6 groups behaved in very similar, linear-like patterns: as the passage difficulty increased, their second-pass dwell times also increased. In contrast, Level 2 group showed a sharp increase between the intermediate- and the advanced-level passages.

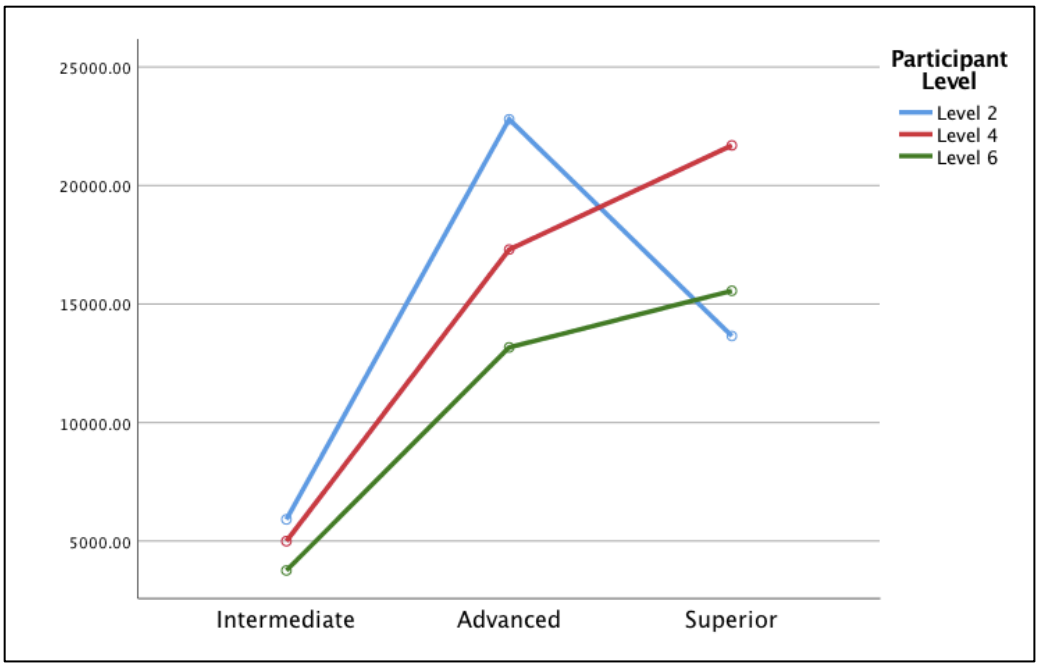


Figure 7. Text second-pass dwell time changes across passage level by student level.

Stem second pass dwell time.

We examined the amount of time in milliseconds participants read the entire stem the second time. This examination included all fixations across the second run as students interacted with the stem only (see Figure 7). The analysis shows that student level $F(2, 47) = .044$, $p = .957$ was not significant but the text level $F(2, 94) = 4.416$, $p = .016$ was statistically significant. The interaction effect between student level and text difficulty was not statistically significant, $F(4, 94) = .917$, $p = .454$ The collected results for student level by text difficulty are displayed in Table 6.

Table 6. *Average Stem Second-Pass Dwell Time for Student and Difficulty Levels in Milliseconds*

		Text Difficulty						
		N	Intermediate	SD	Advanced	SD	Superior	SD
Student Level	Level 2	14	542.43	442.404	950.29	872.852	485.21	451.296
	Level 4	17	867.76	1,038.250	731.53	822.371	449.29	341.175
	Level 6	19	637.95	691.397	935.63	890.147	557.79	427.930
	Total	50	689.34	773.144	870.34	850.997	500.58	401.948

The data revealed the following patterns. First, it appears that participants needed less time to read the question stems the second time around. For example, readers spent around 1.4 seconds reading the advanced-level question stems the first time. In contrast, the second-pass time only took about 870 milliseconds. Another pattern has to do with the fact that the superior-level question stems were read faster than the others. In terms of student level, it appears that the Level 2 and Level 6 groups behaved in similar ways. Both groups showed a sharp increase between the intermediate and advanced-level stems before a sharp decrease while interacting with the superior-level stems during the second pass.

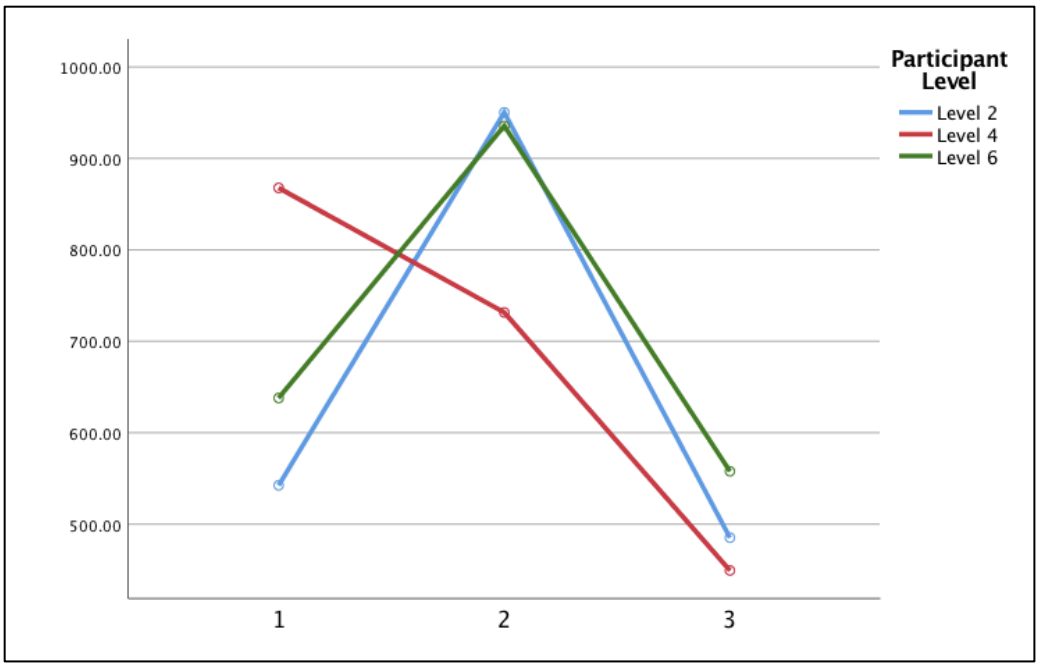


Figure 8. Stem second pass dwell time changes across passage level by student level.

In sum, Table 7 indicates areas where careful and expeditious reading measures proved to be significant. As a reminder, look-backs measure expeditious reading and dwell times measure careful reading.

Table 7. *Summary of Results*

Significant and Non-Significant Results			
	Student level	Text level	Interaction
Passage Look-backs	p=.548	p=.282	p <.001
Stem Look-backs	p=.369	p=1.47	p= .476
Passage 1st Pass Dwell Time	p= .407	p= .007	p= .699
Stem 1st Pass Dwell Time	p= .410	p= .047	p= .620
Passage 2nd Pass Dwell Time	p= .293	p < .001	p= .301
Stem 2nd Pass Dwell Time	p= .957	p= .016	p= .454

Discussion

Despite some of our results not being statistically significant, we noticed fascinating differences in the way readers from varying levels interact with texts. This study collected eye-movement data (first- and second-pass dwell time and look-backs) on the reading passage and question stem of eight reading comprehension items. These eight items were further divided into three groups. The levels of text difficulty were intermediate, advanced, and superior. In this section, we discuss our interpretation of the differences that emerged.

Passage Look-Backs

In accordance with Weir's (2009) reading model, we chose to collect look-backs data in this study. Look-backs are a measure of expeditious reading, and we wanted to know how our three groups of readers varied within this measure. As we interpret our results, it is important to keep in mind that more look-backs indicate a more expeditious type of reading.

When we take a look at Level 2, we notice a trend. This group made more look-backs when reading at the intermediate level, and as they moved on to advanced-level passages, look-backs decreased. It is important to mention that as these readers engaged with the superior-level passages, more look-backs were made than with the previous difficulty level. Since look-backs measure expeditious reading, we can then assume that Level 2 readers attempted to use more strategies while reading the passages that were closer to their level. The decrease in look-backs with the upper levels could indicate being overwhelmed with the difficulty of the text. The Level 4 group made the most look-backs when reading the passages at the advanced level. Once again, this may be an indication of expeditious reading. What is interesting about this group is that when reading the intermediate and superior passages, they behaved similarly. In both cases, look-backs were fewer than when reading the advanced passages. As we attempt to interpret

these results, we can conclude that fewer look-backs were needed while reading the intermediate passages because of how easy they were. As far as the look-backs made while interacting with the superior passages, we could assume that this was due to the difficulty of the text. Finally, Level 6 made the most look-backs when engaged with the passages at the superior level. As the passage difficulty increased, Level 6 readers made more look-backs. Although there was a numeric pattern, students across all three proficiency levels failed to exhibit differences in their expeditious reading processes irrespective of passage difficulty level. However, since the interaction was significant, this indicates that student level and difficulty were not contributing factors on their own but were together. In other words, certain combinations of student level by text difficulty were significantly different than other combinations. Based on Table 1, it seems that Level 2 students reading intermediate texts were very different from the other two levels at intermediate; also, everyone was fairly similar at the advanced level. Furthermore, Level 4 at superior was very different from the other two level groups. So as Level 2 students progressed in reading text difficulty, their expeditious reading decreased, as did Level 4 when they approached superior texts. Overall, data suggests that students read less expeditiously on texts that are above their reading level.

Stem Look-Backs

Before interpreting the results from the question stem look-backs data, it is important to mention that a look-back occurs when a subject returns to an AOI after previously leaving it (Holmqvist, 2011). We operationalize look-backs as a measure of expeditious reading, so the more look-backs a reader makes, the more strategic he or she can be considered while reading.

Level 2 and Level 4 groups behaved in similar patterns. Both groups made the fewest look-backs when reading at the advanced level. The major difference between these two groups

is that Level 2 made more look-backs across all three difficulty levels. Finally, Level 6 behaved as we expected. They made more look-backs as the question stem difficulty increased. This pattern tells us that they were possibly more strategic in their quest to answer the comprehension questions correctly.

The results from this study appear to indicate that when readers engage with test items of varying difficulties at the global level, there are no differences in how expeditiously they read. The three groups did not provide evidence of differences in their expeditious reading of question stems across the three levels of difficulty.

Passage First Pass Dwell Time

One of the first observations that stands out from this set of data is the reading pattern exhibited by the Level 4 group. Our data indicates that this group's first-pass dwell time was very linear. As the passage difficulty increased, so did the amount of time they read the passage for the first time. In fact, all three groups show this pattern. The main difference is that Level 2 and Level 6 groups mirrored one another: dwell times decreased as they read the superior-level texts. This study uses first-pass dwell time as proxy for careful reading: the longer the dwell times, the more careful reading taking place. We can conclude that the intermediate-level passages required little careful reading, possibly due to their word length and content. We can assume that the advanced-level passages were read more carefully, that is, slowly, carefully, linearly, and incrementally, in order to extract meaning from the text (Khalifa & Weir, 2009). From a mathematical perspective, there is no difference in the way L2 readers carefully read the different level passages based on student levels. However, statistically, text difficulty may affect student's usage of careful reading. It appears that when reading at the global level, readers

process difficult texts more carefully. Overall, this study shows that as text difficulty increases, readers tend to read more carefully regardless of the students' level.

Stem First-Pass Dwell Time

Once again, in our interpretation of results, dwell time is a measure of careful reading (Weir et al., 2009). With this in mind we can conclude that it appears as if Level 2 and Level 6 read the stems in similar ways. Both groups dwelled on the advanced stems essentially the same amount of time. As with the passage first-pass dwell time, these two groups mirrored one another once again. After spending the most time reading the advanced-level stems, dwell time plummeted as the superior-level stems were read. Our data analysis appears to indicate that this may not be caused by the student level but by text difficulty as this variable was found to be statistically significant. Regardless of their level, students appeared to read more carefully as the stem difficulty increased. In other words, the participants in this study had a relatively good understanding of what careful reading at the global level looks like.

Passage Second-Pass Dwell Time

Our data indicates that the three groups of readers behaved in somewhat similar patterns as they read the passages the second time. All three groups spent the least amount of time carefully reading the intermediate-level passages. Not only that, but perhaps as expected, the Level 6 group spent the least amount of time dwelling, followed by the Level 4 group, then the Level 2 group. The same pattern is true for the advanced-level passages; the only difference is that the Level 2 group had a much sharper increase in dwell time compared to the other groups. Our interpretation is that text difficulty can affect how readers from various levels read carefully. This interpretation is supported by our statistical analysis, which indicates that text difficulty may be a factor in how students choose to read carefully. The main finding here is that the

reader's level does not affect how he or she reads carefully at the global level. All of our groups showed their ability to read carefully as the difficulty of the text increased. This result may be because these readers have had experience with careful reading even at the lower levels.

Stem Second-Pass Dwell Time

It appears that Level 2 and Level 6 readers behaved in similar patterns as they read the question stems for the second time. Both of these groups dwelled the longest on the advanced-level stems. This dwell time suggest an effort to read slowly and carefully. All of our three groups spent the least amount of time reading the superior-level stems. Our data analysis indicates that student level is not a factor in how carefully readers interact with stems the second time they read them. However, our data suggests that text difficulty may be a contributing factor in how carefully students read question stems at the global level when these vary in difficulty. In other words, all three groups of readers demonstrated their ability to adapt to the difficulty of the text: as the difficulty increased, so did their careful reading patterns.

This study investigated the reading behaviors of English-language learners from varying levels on reading assessment tasks from different difficulty levels. The results from the ANOVAs indicate that there are no statistically significant differences based on student level. In other words, our three groups of readers behaved more or less in similar fashion. Despite these results, interesting differences among the three proficiency groups emerged. First of all, it appears that when reading text passages, the Level 6 group made more look-backs as the passage difficulty increased. In other words, this group of readers demonstrated their ability to adapt to text difficulty and read expeditiously. The opposite was true for the Level 2 group. As text difficulty increased, look-backs, or expeditious reading, also declined. Our question stem data revealed a similar pattern. Level 6 readers, in a linear fashion, made more look-backs as stem

difficulty increased. In terms of first-pass dwell time while interacting with the reading passages, the Level 2 and Level 6 groups behaved in similar patterns. Both groups of readers dwelled longer on the passages as the difficulty increased with the exception of the superior level. Level 4 group read stems in a very linear fashion, and we see this as an indication of careful reading. As the three groups interacted with the question stems, group Levels 2 and 6 behaved similarly. Second-pass time dwell time data showed that our three groups read the passages in similar ways, dwelling longer as difficulty increased. The stem data suggests a similar pattern. Overall, it appears that student level does not affect readers' ability to access expeditious and careful processes. However, text difficulty may affect how long students dwell on passages and stems and therefore affect careful reading.

Conclusion

Summary

The present study used look-backs as proxy for expeditious reading and first- and second-pass dwell time as measurements of careful reading. Fifty ESL students participated in the study, and their eye-movements were recorded as they read and responded to reading assessment items.

The results show that in terms of expeditious reading at the global level, it is not simply the student's level or the text's level that have an impact on strategic reading, but rather, these two variables interact with one another. In other words, variability in expeditious reading appears to be determined by the interaction of student and text level. As far as careful reading, data suggests that text difficulty may be a factor in how readers engage with the text. Our data shows that regardless of student level, our groups of readers were able to adapt to text difficulty and read more carefully as the difficulty increased. This result was true for both the passages and the stems. It appears that these readers have a relatively good understanding of how to read carefully.

Limitations

Despite our greatest efforts, there are some limitations in the study. First, there were only 9 test items, and one had to be removed. A wider range of test items may allow us to more clearly see the interaction between student level and text difficulty. Second, we did not collect data on the participants' responses to the multiple-choice questions. Being able to collect data on the correct answer choice and the distractors for each test item would have allowed us to dissect how effective careful and expeditious reading were.

Future Research

Future research can focus on addressing why students did not differ significantly in their reading behaviors. It would be wise to collect data from the distractors and correct choice for each test item. Also, future studies could look into the differences between successful and unsuccessful test takers. It would be interesting to examine exactly how effective careful and expeditious reading truly are.

Implications

An implication from this study appears to indicate that a student's level may not necessarily negatively affect performance on reading tasks; however, text difficulty may have an impact. This study applies directly to classroom instruction. Our results indicate that ESL students are able to read carefully and that their level does not prevent them from using this type of reading. However, more time and instruction should be devoted to expeditious reading strategies. Teachers should make time to teach their students how to skim, scan, and word search, as all these are components of expeditious reading.

References

Allan, A. (1992). Development and validation of a scale to measure test-wiseness in EFL/ESL reading test takers. *Language Testing*, *9*(2), 101–119. https://doi.org/10.1177/026553229200900201

Anderson, N. J. (2003). Scrolling, clicking, and reading English: Online reading strategies in a second/foreign language. *The Reading Matrix*, *3*(3), 1–15.

Anderson, N. J. (2017). Individual differences in strategy use in second language reading and testing. *The Modern Language Journal*, *75*(4), 460–472.

Ashby, J., Yang, J., Evans, K. H. C., & Rayner, K. (2012). Eye movements and the perceptual span in silent and oral reading. *Attention, Perception & Psychophysics, 74*(4), 634–640. Retrieved from https://search.proquest.com/docview/1323346254?accountid=4488

Aslanian-Nemagerdi, Y. (1986). *An investigation of the reading strategies of five proficient and five less proficient advanced ESL students* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Global. (ProQuest Document ID No. 303464717)

Bahmani, R., & Mohammad, T. F. (2017). Effects of different text difficulty levels on EFL learners' foreign language reading anxiety and reading comprehension. *Reading in a Foreign Language, 29*(2), 185–202. Retrieved from https://search.proquest.com/docview/1979192802?accountid=4488

Baleghizadeh, S., & Yousefian, M. (2012). The relationship between test-takers' rate of strategy use and their reading comprehension performance. *Social Science Premium Collection, 48*(1), 73–115.

Bax, S. (2013). The cognitive processing of candidates during reading tests: Evidence from eye-tracking. *Language Testing*, *30*(4), 441–465. https://doi.org/10.1177/0265532212473244

Bax, S., & Weir, C. (2012). Investigating figurative proficiency at different levels of second language writing. *Research Notes*, *47*(47), 14–28.

Bertram, R. (2011). Eye movements and morphological processing in reading. *The Mental Lexicon, 6*(1), 83–109. Retrieved from https://search.proquest.com/docview/881461334?accountid=4488

Beymer, D., Orton, P. Z., & Russell, D. M. (2007). An eye tracking study of how pictures influence online reading. In C. Baranauskas, P. Palanque, J. Abascal, & S. D. J. Barbosa (Eds.), *Human-Computer Interaction—INTERACT 2007* (pp. 456–460). Berlin, Heidelberg: Springer Berlin.

Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Current Directions in Psychological Science*, *23*(3), 201–207. https://doi.org/10.1177/0963721414530145

Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, *39*(3), 206–219. https://doi.org/10.1016/j.cedpsych.2014.05.003

Brunfaut, T., & Mccray, G. (2015). *Looking into test-takers' cognitive processes whilst completing reading tasks: A mixed-method eye-tracking and stimulated recall study*. *ARAGs Research Reports Online. AR, 2015/001*. London: British Council.

Chiang, M. (2016). Effects of varying text difficulty levels on second language (L2) reading attitudes and reading comprehension. *Journal of Research in Reading*, *39*(4), 448–468. doi:10.1111/1467-9817.12049

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 341–371).

Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-tracking: A guide for applied linguistics research*. Cambridge University Press.

Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PLoS One, 10*(8). doi.org/10.1371/journal.pone.0134008

Gibb, B. G. (1964). *Test-wiseness as secondary cue response*. Retrieved from ProQuest Dissertations & Theses Global (302180597). Retrieved from https://search.proquest.com/docview/302180597?accountid=4488

Goodman, K.S. (1972). The reading process: Theory and practice. In R. Hodges & E. Rudorf (Eds.), *Language and learning to read.* Boston: Houghton-Mifflin.

Gough, P.B. (1972). One second of reading. In J. Kavanagh & I. Mattingly (Eds.), *Language by ear and eye* (pp. 331–358). Cambridge, MA: MIT Press.

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

Grabe, W., & Stoller, F. (2002). *Teaching and researching reading.* New York: Longman.

Grellet, F. (1981). *Developing reading skills.* Cambridge, England: Cambridge University Press.

Guo, Q., Kim, Y. –S. G., Yang, L., & Liu, L. (2016). Does previewing answer choice options improve performance on reading tests? *Reading and Writing*, *29*(4), 745–760. https://doi.org/10.1007/s11145-016-9626-z

Harasym, P. H., Price, P. G., Brant, R., Violato, C., & Lorscheider, F. L. (1992). Evaluation of negation in stems of multiple-choice items. *Evaluation & the Health Professions*, *15*(2), 198–220. https://doi.org/10.1177/016327879201500205

Hedgcock, J., & Ferris, D. R. (2009). *Teaching readers of English: Students, texts, and contexts*. New York, NY: Routledge.

Holmqvist, K. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford, UK: Oxford University Press.

Hyönä, J., Lorch, R. F., & Rinck, M. (2003) Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research (*pp. 313–334). Oxford, UK: Elsevier Science.

Just, M. A, & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. https://doi.org/10.1037/0033-295X.87.4.329

Just, M., & Carpenter, P. (1987). *The psychology of reading and language comprehension.* Boston, MA: Allyn & Bacon.

Katalayi, G. B., & Sivasubramaniam, S. (2013). Careful reading versus expeditious reading: Investigating the construct validity of a multiple-choice reading test. *Theory and Practice in Language Studies*, *3*(6), 877–884. https://doi.org/10.4304/tpls.3.6.877-884

Khalifa, H., & Weir, C. J. (2009). Examining reading: Research and practice in assessing second language reading. *Studies in Language Testing 29,* Cambridge, UK: Cambridge University Press.

Koda, K. (2004). *Insights into second language reading: A cross-linguistic approach.* Cambridge, UK: Cambridge University Press.

Lee, J. (2011, December). Second language reading topic familiarity and test score: Test-taking strategies for multiple-choice comprehension questions. UMI No. 3494064.

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. M. Underwood (Ed.), *Eye guidance in reading and scene perception.* Oxford, UK: Elsevier.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive demands of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 565–572.

Macalister, J. (2010). Speed reading courses and their effect on reading authentic texts: A preliminary investigation. *Reading in a Foreign Language, 22*(1), 104–116. Retrieved from https://search.proquest.com/docview/1705667104?accountid=4488

McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, *19*(2), 113–139. doi.org/10.1007/s10648-006-9010-7

McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology*, *35*(4), 229–241. https://doi.org/10.1016/j.cedpsych.2009.12.001

Nation, P. (2009). Reading faster. *International Journal of English Studies, 9*(2), 131–144, 168–169. Retrieved from https://search.proquest.com/docview/236968159?accountid=4488

Park, G.-P. (1997). Language learning strategies and English proficiency in Korean university students. *Foreign Language Annals*, *30*(2), 211–221. doi.org/10.1111/j.1944-9720.1997.tb02343.x

Park, H. (2005). *Second language reading: The interrelationships among text adjuncts, students' proficiency levels and reading strategies* (Order No. 3192443). Retrieved from ProQuest Dissertations & Theses Global. (304994776), https://search.proquest.com/docview/304994776?accountid=4488

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. Retrieved from https://search.proquest.com/docview/70115641?accountid=4488

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Rupp, A. a, Ferne, T., & Choi, H. (2006). How assessing reading comprehension with multiple-choice questions shapes the construct: a cognitive processing perspective. *Language Testing*, *23*(4), 441–474. https://doi.org/10.1191/0265532206lt337oa

Sabatini, J., Bruce, K., Steinberg, J., & Weeks, J. (2015). SARA reading components tests, RISE forms: Technical adequacy and test design. ETS Research Report Series, 2015(2), 1–20. doi:10.1002/ets2.12076

Sternberg, R. J., & Frensch, P. A. (2014). *Complex problem solving: Principles and mechanisms*. New York, NY: Psychology Press.

Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology, 8,* 323–332.

Thorndike, R. L. (1973–1974). Reading as reasoning. *Reading Research Quarterly, 2,* 135–147.

Tran, T. N. Y., & Nation, P. (2014). Reading speed improvement in a speed reading course and its effect on language memory span. *Electronic Journal of Foreign Language Teaching, 11*(1), 5–20. Retrieved from https://search.proquest.com/docview/1718055821?accountid=4488

Wang, Z., Sabatini, J., O'Reilly, T., & Feng, G. (2017). How individual differences interact with task demands in text processing. *Scientific Studies of Reading*, *21*(2), 165–178. doi.org/10.1080/10888438.2016.1276184

Weir, C. J., Hawkey, R., Green, A., & Devi, S. (2009). *The cognitive processes underlying the academic reading construct as measured by IELTS Authors*. *IELTS Research Reports Volume 9* (Vol. 9). Retrieved from http://www.ielts.org/researchers/research/volume_9.aspx

Weir, C. J., & Khalifa, H. (2008). A cognitive processing approach towards defining reading comprehension, Cambridge ESOL: Research Notes, 31, 2–10.

Wolf, M. (2007). *Proust and the squid: The story and science of the reading brain.* New York, NY: HarperCollins.

Yanagawa, K., & Green, A. (2008). To show or not to show: The effects of item stems and answer options on performance on a multiple-choice listening comprehension test. *System*, *36*(1), 107–122. doi.org/https://doi.org/10.1016/j.system.2007.12.003

Yang, P. (2000). *Effects of test-wiseness upon performance on the test of English as a foreign language* (Order No. NQ59700). Retrieved from ProQuest Dissertations & Theses Global, https://search.proquest.com/docview/304659398?accountid=4488