Metacognitive Skills and the Role of Text-Marking in Reading Comprehension

Jiaan Shang

A Thesis

Presented to the Faculty of Mount Holyoke College

in partial fulfillment of the requirements

for the degree of Bachelor of Arts

with Honor

Department of Psychology and Education

Mount Holyoke College

South Hadley, Massachusetts

Acknowledgements

The idea of doing a thesis in my senior year first emerged in summer 2020. I stayed on a very empty campus that summer after COVID-19 hit, working remotely for our lab, with the hope that the pandemic would end very soon. More than a year has gone by, and this project has been one of the things that I look forward to the most in the midst of a pandemic. I truly enjoyed every moment working on it, as if the chaos in other parts of life temporarily faded away. I felt supported and cared for by many people in this process who are the highlights in my college career. (Pun intended)

First and foremost, I would like to thank my thesis advisor, Professor Kathy Binder. Thank you for always believing in me and devoting so much of your time into this project. Every week when we wrap up our meetings, Kathy always says: “This is getting exciting!” You are right it is! You have taught me to review information critically, to ask important questions, and to be comfortable with unforeseen problems in the research process. Your enthusiasm for research in psychology is inspiring, and I feel truly grateful for having you as a mentor at Mount Holyoke!

Our lab manager, Corrin Moss, has contributed greatly to this project. Thank you for always being the first to help when I run into difficulties, whether it being an issue with our data log or that I have trouble translating my thoughts into words. I feel very supported knowing that you are with me every step of the way!

I am extremely grateful for the support I have received from our lab members at the EXPO-RC Lab- Sonali, Brynn, Jingyi, Mary, Alejandra, Rosa, Zoe and Jamie. They have helped me test-run the experiment, collected, entered and coded numerous data, and gave me candid advice on how to improve the study. You are the best lab mates anyone can ask for!

I have received valuable feedback on this project from my academic advisor and thesis committee member, Professor Mara Breen, as well as my honors committee member, Professor Amy Nussbaum at the department of mathematics & statistics. I appreciate your interest in my project, and your feedback has continuously challenged me to think deeper.

Thank you to Cheryl Lavigne, Janelle Gagnon, and Kat Tremblay for helping me promote the study and recruit participants. I would like to extend my gratitude to other faculty members and staff at the department of psychology & education. It has been a pleasure to learn from you in the past few years. You have not only inspired my interest in psychology and research, but also made Reese a fun and close community.

Thank you to all the participants of our study for the time and energy you have dedicated. This project would not have been made possible without you.

Finally, I want to acknowledge my family and friends for always having my back. My greatest gratitude goes to my parents and my grandmother, who encouraged me to pursue my passion and supported all of my decisions. Even though we are an ocean apart, hearing your voices over the phone for a few minutes each day gives me all the courage I need to persist. I have also received lots of encouragement and inspirations from my friends at and outside of Mount Holyoke. You are the people whom I shared the joy, sadness, frustration and pride with. As my undergraduate studies have come to an end, I find it important to share this final moment with you as well.

Abstract

Text-marking is a popular reading strategy among students, but there has been much debate in empirical literature on its effectiveness. The current study examined the relationship between readers' metacognitive skills and the quality of their text-marking, as well as the impact of text-marking on reading time, question answering time and comprehension accuracy. Specifically, we asked if the effect of text-marking relied on the physical act of marking or seeing the marks in text. 90 undergraduate students were separated into three groups and completed a reading task with six passages. One group was provided with a highlighter and could access their highlighted text when answering questions, one group could highlight but answered questions with a clean text, and the other group did not use highlighting. Readers who were provided with a highlighter spent significantly more time reading the text, regardless of how many words they actually highlighted, but this increase in reading time did not translate into any benefits when answering questions. Additionally, contrary to our hypothesis, readers with higher metacognitive skills highlighted more words and less central words. We recommended that students should use text-marking with caution and instructors should inform students of the cognitive processes involved in text-marking.

TABLE OF CONTENTS

Acknowledgements ii

Abstract iv

TABLE OF CONTENTS v

LIST OF FIGURES vi

LIST OF TABLES vii

Introduction 1
- College Students' Reading Comprehension 1
- Metacognitive Skills 3
- Effect of Text-Marking 7
- Cognitive Mechanism Behind Text-Marking 11
- Current Study 13

Methods 16
- Participants 16
- Materials 17
- Procedure 19
- Analysis of Results 20

Results 22
- Metacognitive Skills and Quality of Text-marking 22
- Reading Time 25
- Accuracy and Question Answering Time 28
- Quality of Text-Marking and Search Efficiency 31

Discussion 32
- Limitations 37
- Future Directions 39
- Implications 40

Appendices 43

References 68

LIST OF FIGURES

Page

Figure 1. The relationship between readers' metacognitive score and number of words highlighted. 24

Figure 2. The relationship between readers' metacognitive score and the centrality of highlighted material 25

Figure 3. Reading time across the three conditions. 27

Figure 4. The relationship between metacognitive scores and reading time. 28

Figure 5. Accuracy across three conditions. 29

Figure 6. Question answering time across three conditions. 30

Figure 7. The relationship between the number of words highlighted and question answering time. 31

Figure 8. The relationship between the centrality of highlighted material and question answering time. 32

LIST OF TABLES

Page

Table 1. The relationship between readers' metacognitive score and number of words highlighted. 23

Table 2. The relationship between readers' metacognitive score and the centrality of highlighted material 24

Table 3. Reading time across the three conditions. 26

Table 4. The relationship between metacognitive scores and reading time. 27

Table 5. Accuracy across three conditions. 29

Table 6. Question answering time across three conditions. 30

Table 7. The relationship between the number of words highlighted and question answering time. 31

Table 8. The relationship between the centrality of highlighted material and question answering time. 32

Introduction

Many students read their class readings or textbooks with a highlighter by hand and came to class with color-coded or underlined materials. The effectiveness of text-marking though, needs to be put under scrutiny through empirical studies. The current study investigated the impact of text-marking on college students' performance on a reading comprehension test, as well as the relationship between readers' characteristics and quality of their text-marking.

College Students' Reading Comprehension

Reading is a skill that students dedicate much time to learn, especially in the early years of elementary school, and as they become more proficient readers, reading itself becomes a critical tool for learning (Goldman et al., 2016). Teachers spend considerable classroom time on reading instructions, and students' reading comprehension skills are assessed regularly as an indicator of their academic competence. In college, reading comprehension is not assessed on its own as often as it is in K-12 education, but it plays a significant role in students' academic success and professional development. The amount of reading required in college is rigorous: According to Gier et al. (2010), college students read about 600-750 pages per class each semester; hence a student taking four classes a semester may read more than 2400 pages each semester. Ryan (2006) argued for the importance of reading in college from two perspectives: Firstly, as instructors cannot cover all materials in class, they expect students to read textbooks and articles before coming to class and engage in class with some understanding of the materials. Secondly, reading textbooks is not only beneficial for understanding the content that the course covers, but improves reading comprehension skills in the discipline in general, which can facilitate future learning. Since reading is such an integral part of students' learning processes, it should not be a surprise that better reading

comprehension skills are related to better academic performance. Specifically, a questionnaire study revealed that students with higher GPA reported more reading goals and more reading strategies compared to those with lower GPA (Taraban et al., 2000). In addition to coursework, reading comprehension is still assessed on the major types of standardized testing required by graduate schools in the United States, such as the GRE, GMAT, LSAT and MCAT. All of these demonstrate the expectations and importance of reading for college students: Students are expected to engage with a larger amount of texts that are more sophisticated and unfamiliar, with precision, critical thinking and insights. Their reading comprehension ability influences their academic performance in college and their professional development after graduation.

Because reading appears to be a skill that students have received many years of training in by the time they enter college, it is natural for instructors to believe that college students should know how to read, and that it is not their role to teach reading strategies to students. In reality, however, not all college students, or even college graduates, are skillful readers. Studies have shown that more and more students have stopped reading class materials, despite many of them recognizing the importance of reading in their learning (Berry et al., 2011; Clump et al., 2004). While this may be due to several factors, including instructor expectation, time management, and lack of interest in the subject matter, one important factor is that students lack the necessary reading skills to help them succeed. Take textbooks as an example. The average length of a chapter is about 50 pages, and it is easy for students to feel overwhelmed by the large amount of information, especially if they cannot separate the important information from other information (Bell & Limber, 2010; Ryan, 2006).

Just because one is not a skillful reader does not mean that they cannot be a strategic reader. Paris et al. (1983) stated that reading strategies are “a skill under consideration”, meaning that

strategies and skills are closely related to each other, except that strategies require more intentionality, and using reading strategies can compensate for the lack of reading skills in some cases. Students use a variety of strategies to read more efficiently, and some of them lack empirical evidence on their effectiveness. Therefore, in order to become strategic readers, it is important for both instructors and students to understand: a) whether a reading strategy is effective and b) how to use it appropriately. The strategy we will focus on in this study is text-marking, specifically highlighting, and we aim to study the cognitive processes of readers whose use of text-marking are successful and unsuccessful. One of the factors that impact readers' strategy use is metacognitive skills.

Metacognitive Skills

Metacognitive skills in reading are a reader's ability to plan, monitor and regulate their reading behaviors (Burin et al., 2020). Readers with good metacognitive skills set their goals before reading, evaluate their comprehension throughout the process, and if they capture inconsistencies or breakdowns in comprehension, they will repair them by replanning and using a different strategy (Zargar et al., 2020). Burin et al. (2020) argued that in studies of reading comprehension, there are multiple dimensions of metacognition, which should not be confused with one another. One type of metacognition is metacognitive knowledge, which refers to students' evaluation of the effectiveness of different learning strategies. Another type is metacognitive evaluation, which requires students to evaluate the difficulty of texts, how much they have understood and how well they will do on a test, etc. The third type is metacognitive strategic activities, which is to employ strategic activities to regulate one's comprehension and fulfill one's reading goals, such as adjusting reading speed and re-reading based on one's monitoring of their comprehension. Metacognitive strategic activity is the focus of the current study.

Multiple studies have demonstrated a significant correlation between metacognitive skills and reading comprehension across age groups. Zargar et al. (2020) investigated comprehension monitoring in third through fifth-grade children by inserting words that were implausible under a given context and using eye-tracking methodology to observe students' reading processes. The authors found that children with stronger comprehension and stronger vocabulary knowledge demonstrated more attempts in comprehension regulation, indicated by longer gaze duration and rereading on implausible words. A similar result occurred for adolescents and college students using a self-reported metacognitive awareness inventory. For high school students, it has been reported that there is a strong, positive relationship between high school students' metacognitive awareness (measured using Mokhtari, Dimitrov and Reichard (2018)'s inventory) and their reading achievement score on a state-level standardized test (Hong-Nam et al., 2014). For college students, those who reported using more metacognitive reading strategies on the inventory (Mokhtari et al., 2018) were more likely to correctly answer reading comprehension questions after reading two expository texts that they had little prior knowledge about (Burin et al., 2020). Apart from self-reported strategy use, think-aloud protocols are often used in studying readers' reading processes as well. Yeari and Lantin (2021) used a think-aloud protocol to investigate the surface-level and deep-level processings of poor and good readers when reading. Surface-level processes refer to activities at the language level such as repetition and paraphrasing, whereas deep-level processes involve connecting different information in the text and making use of prior knowledge. Metacognitive activities are typically considered deep-level processes, as they require careful monitoring of one's comprehension and deliberate allocation of attention. The results in Yeari & Lantin's study showed that good readers used significantly more deep-level thinking like making inferences compared to surface-level thinking, and the proportion of deep-level thinking

is positively correlated to readers' performance. All of these results suggest that the relationship between readers' metacognitive ability and their comprehension performance is significant.

In addition, training on metacognitive skills was shown to be effective in improving comprehension for readers in different age groups (Boulware-Gooden et al., 2007; Fesel et al., 2016; Salmerón & Llorens, 2019). Fesel et al. (2016) taught four metacognitive strategies to 6th grade students: planning (reviewing the appearance of the text), evaluating the importance of information, monitoring their comprehension and predictions, and elaboration. The authors found that while the training group and the control group had a similar pre-test score, those who received metacognitive training scored significantly higher than students in the control group during the post-test. Salmerón and Llorens (2019)'s study instructed similar strategies (planning, evaluating, adapting and monitoring) and demonstrated similar results among 9th grade students. Moreover, they argued that when instructing metacognitive strategies, it was helpful to display eye-movement recordings of readers' who used those strategies successfully and those whose strategy use were unsuccessful, given that some of the mental activities involved in metacognition could be hard to verbalize. Readers in both the eye-movement intervention group and the standard intervention group improved comprehension after the instruction, and those who received instructions with eye-movement models scored significantly higher on a comprehension test than those in the standard intervention group. While there are many limitations to these instructions and the effect they produce, these studies show promising preliminary results that explicit teaching on metacognitive skills are beneficial for students.

Mokhtari et al. (2018) further broke down metacognitive strategies into three major categories, which was the inventory we used in the current study. The first type is global reading strategies, such as having a purpose in mind when reading, previewing the text, and checking if the text fits

one's reading purpose, all of which readers use to set up the stages of reading. These strategies are often included as a part of metacognitive training in intervention studies, as they prevent readers from disorientation and help readers decide when to use other strategies and attend to text that are relevant to their reading goals (Fesel et al., 2016; Salmerón & Llorens, 2019; Van Der Schoot et al., 2008). A correlational study found that among the different reading strategies reported by college readers, setting and monitoring of reading goals was the strategy that differentiated students with high and low performances (Taraban et al., 2000). The second type of strategy is problem-solving strategies like re-reading, adjusting reading speed and getting back on track when distracted. These strategies are more localized, focusing on readers' ability to repair comprehension when a disruption occurs. Using eye-tracking and think-aloud methods, multiple studies have shown that while all readers can detect inconsistencies or errors in texts to some extent, stronger readers are more likely to detect inconsistencies with a larger distance, and they show more attempts to repair their misunderstanding through re-reading (Connor et al., 2015; van der Schoot et al., 2012; Zargar et al., 2020). The third category of metacognitive strategies is supporting reading strategies, where readers make use of external tools to engage with the text. A few examples include taking notes and using reference materials like dictionaries. While all three types of strategies evaluate metacognitive skills, they serve different functions in reading and involve different levels of cognitive processing. Therefore, it remains to be seen whether different types of metacognitive strategies influence reading comprehension to different extent.

It is critical to point out though, that there is a distinction between reading strategies and metacognitive strategies. Indeed, reading strategies are characterized as a deliberate, goal-directed and controlled act, requiring readers to be metacognitive and understand their reading goals and processes (Afflerbach & Cho, 2009; Afflerbach et al., 2008). However, these strategies are

sometimes not used in combination with metacognition in reality. For example, both notetaking and text-marking are popular strategies used by students, but they can easily become a mindless physical act without enough processing and comprehension monitoring. As Yeari et al. (2017) mentioned in a study on the effects of highlighting, highlighting is an external move, whereas differentiating central information is a mental process, and while the two processes are often intertwined, they can yield different products depending on the mental processes each reader engages in while they highlight. A similar argument has been raised by multiple researchers that it is not the reading strategies themselves, but rather the way that they are implemented, that determines whether they improve comprehension and recall or not (Dunlosky et al., 2013). Effective use of reading strategies brings readers closer to their goals whereas inappropriate strategies hinder their comprehension. Meanwhile, there has been limited research on the relationship between readers' characteristics, namely metacognitive skills, and how effectively students use reading strategies such as text-marking. Therefore, this study aims to bridge this gap and investigate whether the effectiveness of text-marking is influenced by readers' metacognitive skills.

Effect of Text-Marking

Text-marking, including highlighting and underlining, is one of the most common learning strategies among students (Yue et al., 2015). There are typically two types of text-marking: pre-existing text-marking and reader-generated text-marking. Pre-existing text-marking refers to both typographical cues such as bolding and italicizing by publishers, and markings made by previous textbook users. Multiple studies showed that the pre-existing markings can be beneficial for both comprehension and retention, but only when an appropriate amount of texts were marked and when the marked sections were relevant (Gier et al., 2009; Lorch et al., 1995). In a study conducted

by Lorch et al. (1995), college students were asked to read a 4-page text with light signaling, heavy signaling or no signaling. In the light signaling condition, 14 target sentences were marked, which made up about 5% of the text, whereas in the heavy signaling condition, 14 target sentences were marked in addition to other filler contents, which was a total of 50% of the text. Light signaling facilitated cued recall better than no signaling or heavy signaling, and there was no significant difference between no signaling and heavy signaling. This result showed that text-marking's effects were only significant when they were used selectively. The study of Gier et al. (2009) generated a more concerning result. They found that inappropriate highlighting impaired both comprehension and metacognition accuracy, meaning that readers either overestimated or underestimated how well they comprehended the passage and how well they answered the questions. The author concluded that inappropriate highlighting directs readers' attention away from the most relevant information, hence impeding comprehension and metacognition.

As the effect of highlighting is largely influenced by its quality, it is no surprise that reader-generated text-marking also produced mixed results. While there are studies that showed benefits in reader-generated text-markings, the effects are often conditional, depending on the learning condition and text difficulty (Fuson-Newsome & Metzger, 2016; Yue et al., 2015). Yue et al. (2015) found that highlighting significantly improved the retention of highlighted information and did not impair retention of unhighlighted information, but this was the case only for readers in a massed learning condition and not those in a spaced learning condition, showing that highlighting is only effective when readers study the highlighted text immediately after the initial read. Another study used excerpts of psychology textbooks, and demonstrated that highlighting increased reading comprehension test scores for difficult texts but not easy texts (Fuson-Newsome & Metzger, 2016). Additionally, there is no significant difference in the comprehension score between readers who

used dual-colored highlighters and single-colored highlighters. The authors argued that the effects of highlighting only appear when it requires a certain amount of cognitive load to process the texts. More studies, though, showed that reader-generated text-marking can be harmful. A correlational study that examined the study behavior and exam scores of students found that highlighting materials from class was negatively correlated with test scores (Gurung 2010). In another study, Peterson (1991) used a 10,000-word history book chapter to examine the cognitive functions of underlining, and found that underlining did not improve recall, and even had a harmful effect on inferential recall. The author provided two explanations for this result. The first reason was that students did not use underlining discriminately. The second reason was that overreliance on underlined information prevented students from building connections between different parts of the text and processing the text as a whole, resulting in a poorer performance on inferential questions.

Despite the importance of high-quality markings, studies showed that students of all ages have some trouble using this strategy effectively (Peterson, 1991; Rickards & Denner, 1979). One common problem is that students mark too much text (Peterson, 1991). One reason why too much marking may not be helpful for recalling information is that the central information is lost in the excessive markings. It is argued that when the amount of text-marking is unlimited, readers may put less effort into distinguishing the most important information, and process the text less deeply (Dunlosky et al., 2013). While there is no clear evidence of how much marking is considered harmful and how much marking optimizes comprehension, as this number varies greatly depending on the text's length and characteristics, Peterson (1991) considered over 20% of markings as a sign of a lack of discrimination in text-marking. Another common problem is that students had trouble identifying the central information that should be marked, which is a

phenomenon across students of all ages. Rickards and Denner (1979) investigated the effect of underlining and adjunct questions on fifth-grade children's recall of texts. 83% of the statements the children underlined were detail-oriented statements, even when they were asked to underline only one sentence per paragraph, or when they were asked post-questions that guided their attention to central ideas. While the authors argued that adult readers were generally better at identifying central ideas, other studies have shown that adult readers' skills vary as well. In the Peterson (1991) study, most students underlined more than 20% of the text. The author also believed that in order to fully process the text and understand whether a piece of information was central to the passage, readers need to read a section more than once. However, 83% of the students reported that they underlined on their first read through the section. Although many studies argued for the harm of excessive markings and peripheral markings, few of them provided empirical evidence that investigated the relationship between the amount of text-marking and comprehension performance, especially for reader-generated text-markings. Therefore, in this study we aim to put this claim under scrutiny and examine whether or not readers spend more time searching for answers when they mark more words during the initial read.

If high-quality text-marking is indeed critical to comprehension and learning, an association between high reading skills and high-quality text-marking is expected. When examining textbook marking practices of undergraduate students, it was found that low-skill readers depended on text-marking more than high-skill readers (Bell & Limber, 2010). Low-skill readers also tended to mark more texts and marked significantly more irrelevant information compared to high-skill readers. Therefore, the authors recommended that students need to monitor the quality of their highlighting through comparison with the instructor's judgment of relevance, and make good use of typographical cues made by textbook editors. While this study provides valuable information on

the relationship between readers' skill level and their text-markings, reading skills consist of many factors, and we hope to focus on the relationship between metacognitive skills and text-markings in our study.

Cognitive Mechanisms Behind Text-Marking

While many studies have investigated the effects of text-marking, fewer have investigated the cognitive mechanism behind this strategy. The most mentioned mechanism was the von Restorff effect, also known as the isolation effect. Researchers argue that as the highlighted information stands out from the rest of the text, students will spend more time reading it, resulting in better recall (Dunlosky et al., 2013; Nist & Hogrebe, 1987). However, it was not clear whether this effect occurred during the initial read or the restudying phase. To address this question, Peterson (1991) investigated whether underlining served an encoding function, review function, or both. An encoding function refers to a deeper processing during initial reading and underlining, which resembles what Miyatsu et al. (2018) described as a generative function -- the elaborative thinking elicited by the implementation of text-marking. A review function refers to deeper processing during studying or reviewing the material. Similarly, Miyatsu et al. 's article described a storage function of text-marking, meaning that marked texts are easier to identify in later studying.

In Peterson's study, students were randomly separated into three groups and were asked to read a 10,000-word history textbook chapter. The first and second group could mark any text while reading, and the third group could not make any marks. When all three groups returned to review the material after a week, only the first group had access to their markings and the second and third group reviewed with a clean copy. The author did not find apparent benefits from underlining in

either the encoding stage or the reviewing stage, as there was no significant difference between the three groups' recall performance.

While this is the only study design that made a distinction between the encoding function and the reviewing function of highlighting, studies that showed an advantage of reader-generated markings compared to experimenter-generated highlighting provided some valuable evidence for the encoding function. If the function of text-marking is to make targeted information more identifiable later, then the central information highlighted by the experimenter should be just as helpful, or even more helpful, considering that readers are not always able to recognize the central information. The highlighted information should also be more likely to be recalled compared to unhighlighted information. But Rickards and August (1975) reported that readers who were able to make their own markings performed better than all other groups, including readers who reviewed central information highlighted by the experimenter. In addition, readers who were instructed to mark the most central concepts remembered more incidental information even when they were not underlined. Researchers explain that when readers highlight texts, their interaction with the text is active and generative, whereas reading preexisting highlighting is a passive activity (Dunlosky et al., 2013). This explanation is consistent with the finding that more skillful readers are more likely to view reading as an active and integrative process than less skillful readers (Saumell et al., 1999). Therefore, it is possible that readers' mental effort to distinguish central information in the process of highlighting, as opposed to its product, determines the function of highlighting.

Although the above studies examined the function of text-marking in reviewing and recalling, the cognitive function of text-marking in comprehension with text access remains to be seen. Reading comprehension tests, unlike studying for a closed-book exam, do not require readers to

remember the relevant information. Therefore, there is a lower in-text memory demand on readers, but a greater emphasis on knowing the general organization of the text, strategically re-reading and searching for relevant information during re-reads (Schaffner & Schiefele, 2013; Vidal-Abarca, Mañá,& Gil, 2010). Therefore, in this study, we investigated whether text-marking has an influence on students' performance on reading comprehension tests, and whether the influence depends on the physical act of highlighting or the readers' access their highlighted text. We adapted the methods used in Peterson's (1991) study to investigate whether there is a difference in comprehension accuracy and time spent answering questions between readers who have access to their markings and who do not have access.

Current Study

Several studies have shown that the quality of text-marking impacts its effectiveness as a reading strategy. Meanwhile, there has been limited research on the relationship between reader characteristics, especially metacognitive ability, and the quality of text-marking. Therefore, the first goal of this study was to understand the influence of metacognitive skills on the use of text-marking. The second goal was to understand the cognitive function of text-marking, specifically, whether the effect of text-marking depends on the visual cue produced.

In this study, participants will be randomly assigned to one of three groups based on different reading strategies: One group will be asked to highlight information they find important and answer comprehension questions with the highlighting present; Another group will also be asked to highlight when reading but will answer questions using a clean text; The final group will not be able to highlight when reading and will also answer questions with a clean text. After reading the passages and answering the questions, all participants will be given a metacognitive

skills inventory where they are asked about the frequency with which they use metacognitive reading strategies.

The study consists of four major questions. The first question is: Does metacognitive ability influence the quality of text-marking? Specifically, we look at the amount of highlighting and the centrality of highlighted information. We hypothesize that participants with higher metacognitive skills will highlight less information and will focus on highlighting central information, while participants with lower metacognitive skills will highlight more information, especially more peripheral information.

The second question is: Do readers spend more time reading the passage while they are highlighting? Previous literature has used reading time as an operational measure of how much attention has been allocated to a certain segment of the text. (Reynolds & Anderson, 1982; Broughton et al., 2010). Specifically, Broughton et al. (2010) argued that cognitive engagement with a text can be separated into online engagement and offline engagement. Offline engagement refers to how information is represented mentally after reading, and can be reflected by readers' performance on comprehension questions, while online engagement refers to readers' engagement with the text while they are reading, and can be demonstrated through reading time. Therefore, it is reasonable to establish that readers spend more time reading the passage when they process the sentences more thoroughly and recognize a sentence's importance in the passage before they highlight. Taking this a step further, we hope to use reading time to distinguish mindless highlighting and selective highlighting. Our first hypothesis is that readers in the two highlighting conditions will spend more time reading the passage compared to readers who do not highlight. Our second hypothesis is that readers with better metacognitive skills will spend more time reading the passage when highlighting compared to readers with lower metacognitive skills.

The third question is: Does the text-marking behavior and the markings improve readers' comprehension, and if so, does the effect rely on the visual cues or the physical act of text-marking? We will evaluate readers' comprehension from two aspects: Firstly, we will compare the comprehension scores across three groups. If the groups that highlight while reading and answer questions with the highlighting present perform significantly better than the other two groups, this will tell us that highlighting could improve comprehension, and this effect depends on the visual cue produced by highlighting. If participants who highlighted performed significantly better than participants who could not highlight, but there was no difference regardless of the presence of highlighting when answering questions, then this tells us that the physical act of highlighting could be sufficient, and the effect does not rely on the presence of the visual cues. Secondly, we will compare the time it took readers from the three groups to answer questions. If both groups that highlight while reading spend less time answering questions, it is an indication that the act of highlighting may help them locate information faster.

The final question is: Does the amount of text-marking impact how fast readers can locate relevant information when answering questions? Previous literature suggested that overmarking hinders recall (Lorch et al., 1995), since the targeted information is not as distinctive when there are too many markings. In addition, studies have shown that readers tend to slow down when reading marked texts. Therefore, it can take readers longer to search for the information they need among many text-markings. A second part of Lorch et al. (1995)'s study showed that readers slow down when reading signaled content, but their reading time on non-signaled texts are unaffected, regardless of whether other contents are signaled or not. In this part of the study though, the texts are signaled through capitalization as opposed to highlighting or underlining. Another study (Yeari et al., 2017) argued that the relationship between reading time and text-markings does not occur

in the initial read, but in the re-reading phase instead. Although readers tend to devote more time to central information in the initial read, they demonstrate more regressions to peripheral information compared to central information if it is highlighted. This result is particularly meaningful to our research question, because unlike studies that use recall as their criterion tasks, readers in this study have access to the text and will re-read some parts of the text when they search for answers. Based on these results, we hypothesize that a larger amount of text-marking is associated with longer time spent on answering questions.

Methods

Participants

Participants (N=90; age range: 18-47 years old, M_{age}=19.71 years old; 90% female, 7.78% non-binary gender or gender fluid, 1.12% agender; 12.4% Black/African American, 40.4% White/European, 49.4% Asian/Asian American, 5.6% Hispanic/Latino/Latina/Latinx, 3.3% other race and ethnicity) were recruited from undergraduate students enrolled at a liberal arts college in the New England area of the United States. The percentage of students from first-years to seniors were: 44.4%, 23.3%, 25.6% and 6.7%. 9.1% identified themselves as an excellent reader, 38.6% as a good reader, 44.3% as an average reader and 8% a poor reader. Participants were recruited through the psychology department's research website and flyers on campus. Those who signed up through the research website can receive 1.5 research credit for an eligible psychology class. Otherwise, they can enroll in a lottery and have the chance to win a $10 gift card.

Materials

Metacognitive Awareness Inventory

To measure metacognitive skills, we used the Metacognitive Awareness of Reading Strategies Inventory (MARSI) developed by Mokhtari et al. (2018). This measure is a revised version of their original creation in 2002. MARSI is a self-reported questionnaire that asks students about how well they know and how often they use a series of metacognitive strategies, measured on a scale of 1-5 (1=*I have never heard of this strategy before*, 5=*I know this strategy quite well, and I often use it when I read*).

The questionnaire contains descriptions of 15 strategies, divided into three subscales: (1) global reading strategies, (2) problem-solving strategies, and (3) supporting reading strategies. Test results showed this inventory is reliable, valid and invariant across gender or race and ethnical groups. The Cronbach's alpha coefficient for internal validity is .0850, and the correlation coefficient for external convergence is statistically significant (p<.001).

Passage Selection

The passages were selected from the Graduate Record Examination (GRE) practice books, as well as a free access website (https://www.usingenglish.com/comprehension/) that host a large collection of English learning resources, including reading comprehension exercises (Gier et al., 2009; Yeari et al., 2015; Yeari et al., 2017). All passages were expository texts, and the topic of the passages included history, art, geography, theatre, and literature. While some studies have used psychological textbooks or passages as their materials (Bell & Limber, 2010; Fuson-Newsome & Metzger, 2016), we hope to reduce the influence of prior knowledge on comprehension accuracy, since the majority of our participants are current or prospective psychology majors. The average length of the passages we used was 458 words. The questions contained both factual questions and

inferential questions. The passages and questions were printed copies with a font size of 12 and single spaces. The majority of the passages fit on one page and the questions take one to two pages.

To make sure the difficulty of those passages is appropriate for college-level readers, we recruited twelve college students to participate in an online pilot study. Participants were instructed to read ten passages through a Google form and answer five questions after each passage. There was no time limit, but they were instructed to record the time they started the study and the time they finished. The six passages that were used in the main study were selected from these ten passages, and the average percentage of correct responses of them was 58.5%, ranging from 30% to 90%. This suggested that the passage is able to differentiate readers of different skill levels.

Centrality Index

In this study, "centrality" refers to an information unit's importance in the passage (McCrudden & Schraw, 2007). The passage was broken down into information units using Yeari et al. (2015) 's method: "Each unit included a main predicate, its arguments (including time and place), and the adjectives and/or adverbs of these arguments". The number of information units for the six passages ranges from 31 to 47, with an average of 35 per passage.

To determine centrality, a group of 10 independent raters rated each information unit in each passage on a scale from 1-5 (1=*least central*, 5=*most central*). An information unit's centrality is defined by two criterias: 1) The information is critical to understanding the main idea of the text 2) Comprehension of the text will be damaged if the information is taken out (Yeari et al., 2017; Yeari et al., 2015). Because most readers highlight on their initial read, especially during a reading comprehension test, we asked that raters give their centrality score on their initial read as well,

although they were allowed to make changes to their ratings based on further information they encountered.

The Cronbach's alpha coefficient for inter-rater reliability of the six passages is 0.82. There are two passages where the inter-rater reliability was below 0.8, where three raters discussed each one of the information units and reached an agreement. The final centrality score for an information is the sum of the ten ratings. Information units with a score above 3.75 (75%) are considered the central information of the text (Miller & Keenan, 2009), which made up about 25% of the information units. The readers' centrality index is the percentage of central words in the total number of words they highlighted. This is used to separate readers who selectively highlight the central information and readers who highlight lots of peripheral information in addition to the central information.

Procedure

After the participants signed the informed consent forms, the experimenter explained the reading task to them and gave them the reading package. They were instructed to read the practice passage and finish the practice questions, after which the experimenter answered questions they had on the procedure, and they moved on to the formal trials.

The formal trials contained six passages, presented in a randomized order. Participants read the passage and were instructed to record the time they began reading and the time they finished reading, using the timer on a lab computer. Participants in the two highlighting groups were provided with a color highlighter and were asked to highlight information while they read as they would normally do in a reading task, and to not make additional markings while they answer questions. For participants who were instructed to highlight while reading but did not have access

to their highlighting later, there was a colored paper between their passage for highlighting and their passage for answering questions to prevent them from seeing their highlighting. Other than that, the reading package was the same across all three conditions. Participants in the no highlighting condition were instructed to not make any marks on the text. After each passage, there were five multiple-choice questions. Participants were asked to record the time they finish answering questions. Participants could answer questions in any order within the same passage and look at the passage, but cannot change answers for previous passages. There was no time limit for this task. Throughout the process, the experimenter stayed in the room to answer questions that the participants may have.

After the reading task, participants were given the Metacognitive Awareness Inventory. They also answered a few demographics questions. Finally, they were debriefed and thanked for their participation.

Analysis of Results

We used linear mixed-effects models to examine all four of our research questions. The significance level of all analyses is α=.05. Outliers in the metacognitive score, reading time, question answering time and the number of words highlighted were detected using the outlier labeling rule: We multiplied the interquartile range (IQR) by a factor of 2.2. The lower bound = 25 percentile - IQR*2.2, and the upper bound=75 percentile + IQR*2.2. Data point outside of this range is considered an outlier and is replaced by either the lowest value within range - 1 or the highest value within range +1.

For the first research question, we investigated whether readers with better metacognitive skills highlight less information and more central information. The readers' MARSI score is the

predictor, and the number of words highlighted is the dependent variable. We hypothesized that readers with a higher MARSI score highlighted less words. Similarly, in our second hypothesis, the MARSI score is the predictor and the centrality index of readers' highlighting is the dependent variable. We hypothesized that readers with a higher MARSI score received a higher centrality index for their highlighting. In both models, the fixed effect is the reader's MARSI score, and the random effects are the participants and passages. An outlier in metacognitive scores that is below the lower bound is winsorized. Outliers in the number of words highlighted that are above the upper bound are also winsorized.

The second research question asked whether or not readers spend more time reading the passage when highlighting, and whether or not readers with higher metacognitive skills will slow down when highlighting. We hypothesized that readers spent more time reading the text when highlighting, and for the two highlighting groups, readers with a higher MARSI score will spend more time reading the passage. Similar to the first research question, the fixed effect is the reader's MARSI score, and the random effects are the participants and passages. Outliers in reading time that are above the upper bound were winsorized.

For the third research question, we compared comprehension performances between participants of the three reading conditions. The two dependent measures were the accuracy score they receive on the reading comprehension test and the time they spend on answering questions. The contrasts will examine differences between participants in the three conditions. Outliers in question answering time that are above the upper bound were winsorized.

For the fourth research question, we want to know whether or not more highlighting leads to a longer time spent on locating information. The dependent variable is the time readers spend

on answering questions. The fixed effect is the proportion of words highlighted, and the random effects are the participants and passages.

Results

Metacognitive Skills and Quality of Text-marking

The quality of text-marking was measured by two dimensions: 1) the number of words highlighted and 2) the percentage of central words highlighted out of all words highlighted, which will be referred to as the centrality index below. The average number of words highlighted is 117.39, which constitute 25.63% of the average number of words in a passage. The average number of central words highlighted is 37.16, and the average centrality index is 34.98%.

Metacognitive score refers to the score readers earned based on the Metacognitive Awareness Inventory. The Cronbach's alpha coefficient for internal reliability of all 15 items is 0.80, and the reliability for each of the three subscales are: 0.57 (global reading strategies), 0.72 (problem-solving strategies) and 0.57 (supporting reading strategies). The average metacognitive score is 62.23 out of 75.

The mixed-effect model showed that there was a significant relationship between readers' metacognitive score and the number of words they highlighted (F=7.543, p=.008). See Figure 1. See Table 1 for coefficients. Contrary to our hypothesis, readers with higher metacognitive scores highlighted more words. The three subscales of the metacognitive inventory, global reading strategies (p=.268), problem-solving strategies(p=.503), supporting reading strategies (p=.243) did not predict the number of words readers highlighted. This relationship is also found when the off-task readers are included in the analyses (F=6.952, p=.011)

The relationship between readers' metacognitive scores and the centrality of their highlighted words was also significant (F=7.500, p=.008), but the direction of this relationship was the opposite of our prediction. Readers with higher metacognitive scores highlighted a lower proportion of central words (See Figure 2. See Table 2 for coefficients). Global reading strategies (p=.151), problem-solving strategies(p=.982), supporting reading strategies (p=.071) alone did not predict the centrality of highlighted material. This relationship was also found when the off-task readers are included in the analyses (F=4.498, p=.038). A follow-up mixed-effects model showed that neither the number of words highlighted (F=0.000, p=.988) nor the centrality index of the highlights are correlated with accuracy (F=0.169, p=.682).

	β	*SE*	*t*	*p*
MARSI Score	3.41	1.24	2.74	.008

Table 1. The relationship between readers' metacognitive score and number of words highlighted.

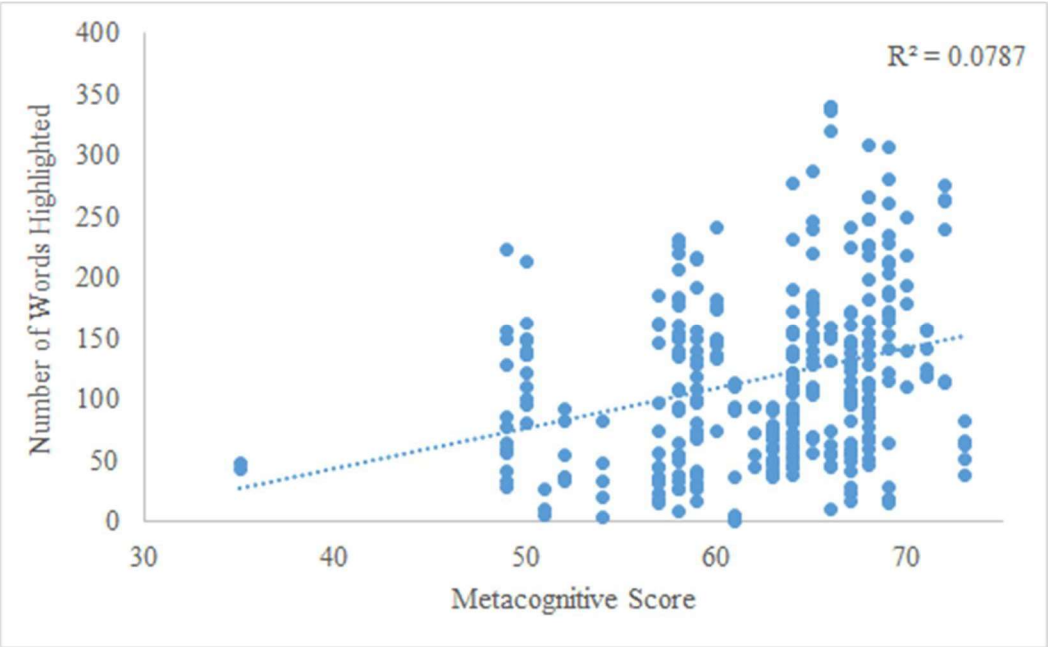

Figure 1. The relationship between readers' metacognitive score and number of words highlighted.

	β	*SE*	*t*	*p*
MARSI Score	-0.46	0.17	-2.74	.008

Table 2. The relationship between readers' metacognitive score and the centrality of highlighted material.

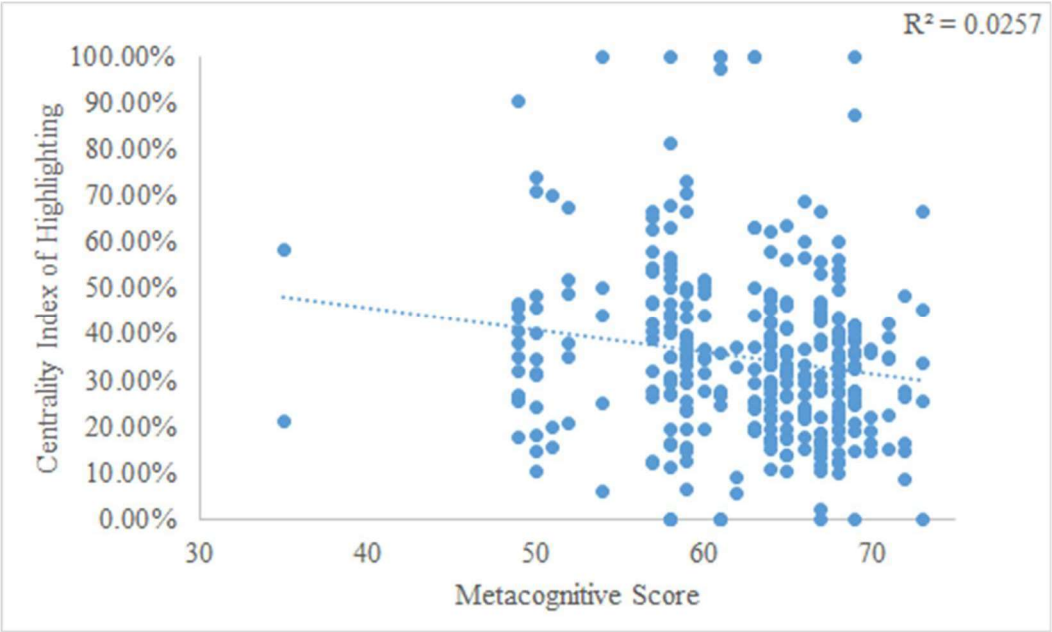

Figure 2. The relationship between readers' metacognitive score and the centrality of highlighted material.

Reading Time

Reading time is reported in seconds. Sixty-five trials were eliminated from the analysis because the readers' reading time was below the minimum time needed to read the passage, which constitutes 12.03% of the total trials. The minimum time needed for reading a passage is calculated by the number of words in the passage times the average time it takes to process a word, which is 250 milliseconds (Rayner & Pollatsek, 1989). If the reading time is lower than this threshold, it is an indication that the reader did not finish reading the passage, hence their trials are eliminated. The average reading time per passage across participants is 215.58 seconds. There is a significant difference in reading time across the three conditions (F=8.310, p<.001). See Figure 3. See Table 3 for coefficients. Readers in the two highlighting conditions spent significantly more time reading the text compared to readers in the no highlighting condition (M=172.539, SD=14.523), but there

is no difference between readers who can access their highlighting later (*M*=234.505, *SD*=14.046) and readers who cannot access their highlighting (*M*=229.646, *SD*=14.565). To examine whether or not this increased reading time is due to the time it takes to physically highlight words, we ran a follow-up mixed-effects model on the relationship between the number of words highlighted and reading time, and found that there is no relationship between these two variables (*p*=.252). This shows that readers spend more time reading the text when they are provided with a highlighter, regardless of how many words they actually highlight.

For readers in the two highlighting groups, there is no significant relationship between their metacognitive scores and their reading time (*F*=2.493, *p*=.120). See Figure 4. See Table 4 for coefficients. Contrary to our hypothesis, readers with higher metacognitive scores did not necessarily slow down when they are asked to highlight.

	β	*SE*	*t*	*p*
Highlighting with Access	61.97	16.67	3.72	<.001
Highlighting without Access	57.11	17.11	3.34	<.001
No Highlighting	0	0	-	-

Table 3. Reading time across the three conditions.

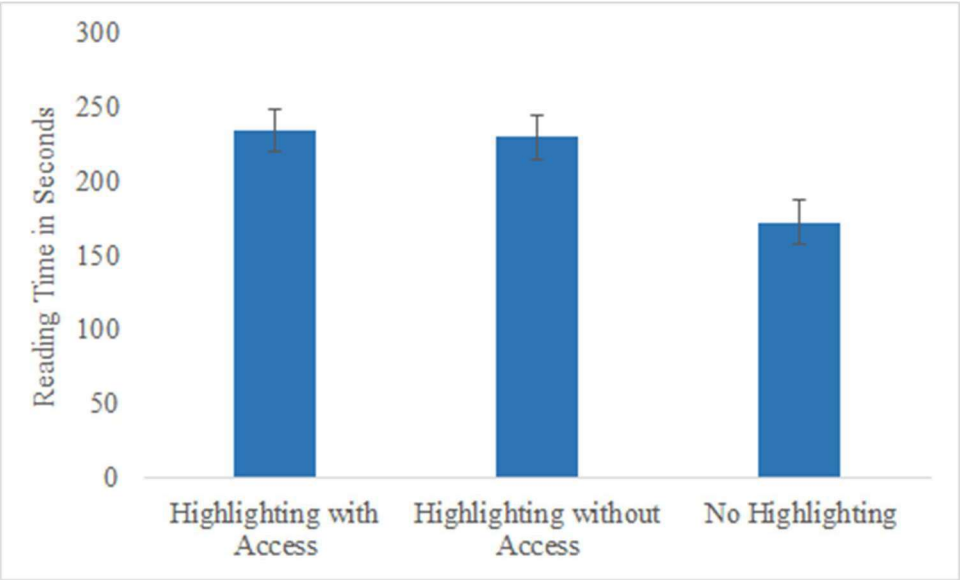

Figure 3. Reading time across the three conditions.

	β	*SE*	*t*	*p*
MARSI Score	1.96	1.24	1.58	.120

Table 4. The relationship between readers' metacognitive score and reading time.

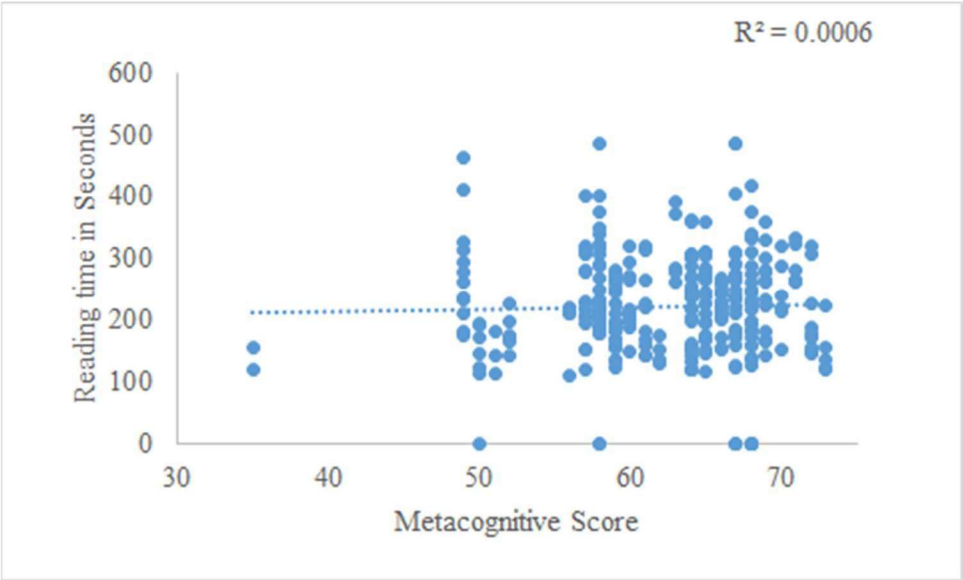

Figure 4. The relationship between metacognitive scores and reading time.

Accuracy and Question Answering Time

Accuracy is measured by the percentage of correct answers out of all questions. Three trials in the two highlighting conditions were eliminated because the reader did not follow the directions and did not highlight any word while reading. The average accuracy across all participants was 59.75%. The accuracy for individual trials ranges from 0% to 100% and the total accuracy for each participant ranges from 16.67% to 90%. There is no significant difference in accuracy (F=0.329, p=.721) between participants who highlighted and had access to their highlighted text (M=60.549, SD=4.769), participants who highlighted and had no access to their highlighted text (M=59.818, SD=4.913) and participants who did not highlight (M=57.037, SD=4.945). See Figure 5. See Table 5 for coefficients. Therefore, we did not find any evidence that either the physical act of highlighting or the visual cues produced by highlighting improved readers' performance on the comprehension test.

	β	*SE*	*t*	*p*
Highlighting with Access	3.51	4.50	0.78	.438
Highlighting without Access	2.78	4.65	0.60	.552
No Highlighting	0	0	-	-

Table 5. Accuracy across the three conditions.

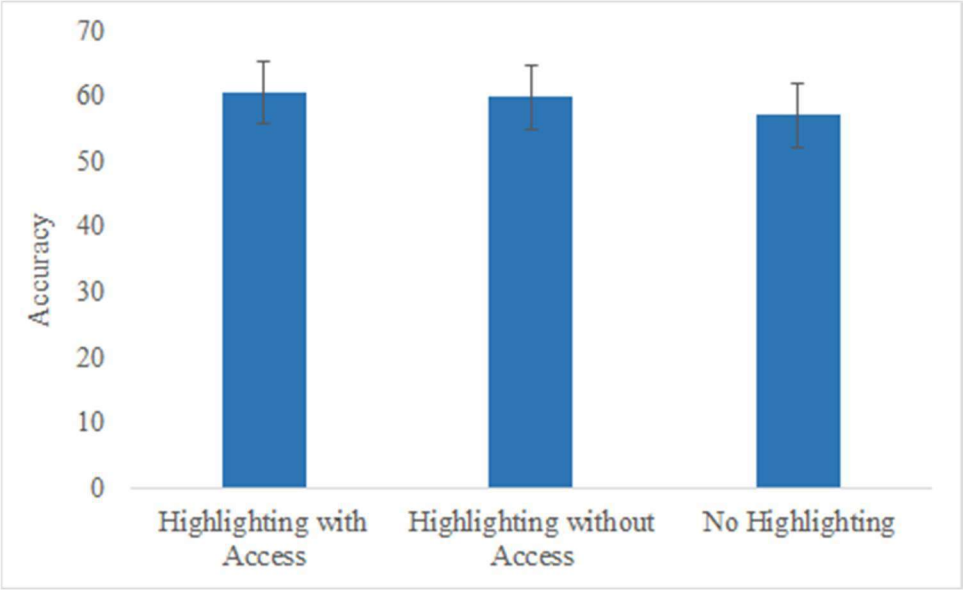

Figure 5. Accuracy across three conditions.

Question answering time is reported in seconds. One trial was eliminated because the readers' question answering time is below the minimum time needed to read the questions using the same method for reading time. On average, readers spend 272.96 seconds per trial answering questions. We did not find a significant difference in question answering time across the three conditions (F=2.427, p=.095). See Figure 6. See Table 6 for coefficients. While there is a trend that the group who have access to their highlighted text when answering questions (M=280.798, SD=24.304) spends more time answering questions than the group who do not have access to their highlighted

text (*M*=240.174, *SD*=24.852) and the group who do not highlight (*M*=238.193, *SD*=24.914), this difference did not reach significance. Overall, there is no significant evidence that highlighting has an effect on the time it takes readers to answer questions. Additionally, follow-up analyses did not find a significant relationship between the time readers spent on answering questions and their accuracy (*F*=1.385, *p*=.240)

	β	*SE*	*t*	*p*
Highlighting with Access	42.60	22.08	1.93	.057
Highlighting without Access	1.98	22.68	0.09	.930
No Highlighting	0	0	-	-

Table 6. Question answering time across the three conditions.

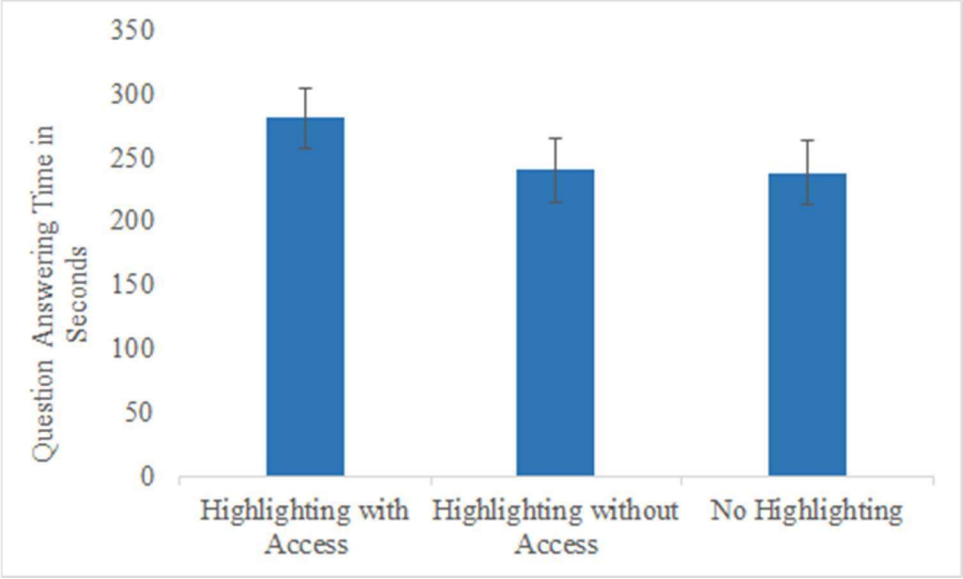

Figure 6. Question answering time across three conditions.

Quality of Text-Marking and Search Efficiency

For participants who have access to their highlighted text when answering questions, there is no significant relationship between the number of words highlighted and the time it takes for readers to answer questions (p=.754). See Figure 7. See Table 7 for coefficients. Contrary to our hypothesis, there is no evidence that excessive highlighting slows down the speed of which readers search for answers in the text. Nor is there a significant relationship between the centrality of the highlighted text and the question answering time (p=.699). See Figure 8. See Table 8 for coefficients. Therefore, there is no evidence that excessive highlighting slows down the speed of searching for answers.

	β	*SE*	*t*	*p*
Number of words highlighted	-0.05	0.17	-0.31	.754

Table 7. The relationship between the number of words highlighted and question answering time.

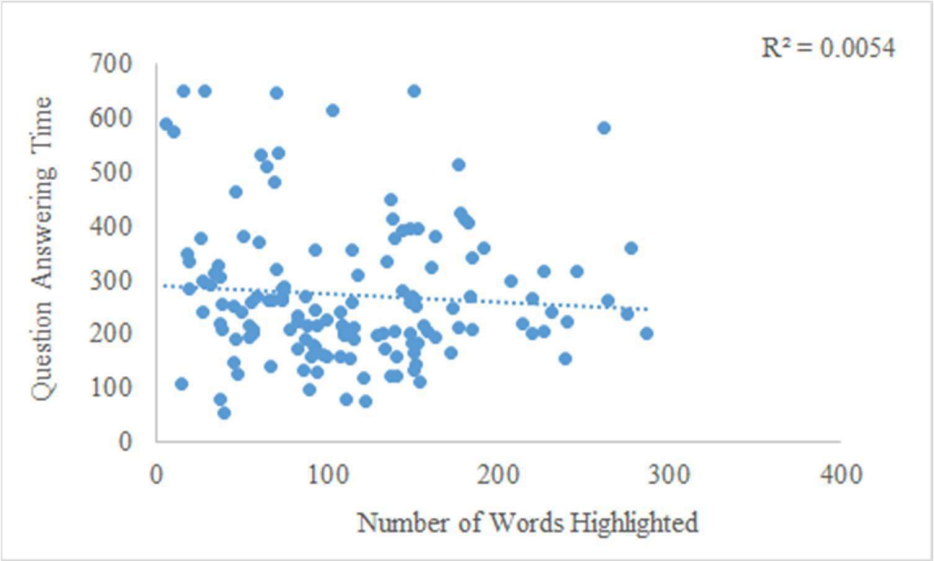

Figure 7. The relationship between the number of words highlighted and question answering time.

	β	*SE*	*t*	*p*
Centrality index	-0.16	0.41	-0.39	.699

Table 8. The relationship between the centrality of highlighted materials and question answering time.

Figure 8. The relationship between the centrality of highlighted materials and question answering time.

Discussion

The goal of this study was to: 1) understand how readers' characteristics, specifically their metacognitive skills, are related to their use of text-marking and 2) examine whether or not text-marking improves comprehension accuracy and efficiency on a reading comprehension test for college students. Overall, we found that readers with higher metacognitive scores highlighted more

words with a lower percentage of central words. Additionally, we did not find significant evidence that text-marking benefits reading comprehension, despite the fact that readers who were provided with a highlighter spent more time reading the text.

Our first research question investigated the impact of metacognitive skills on the quality of text-marking. We are surprised to find that readers with better metacognitive skills produced more peripheral highlighting, considering that readers with better metacognitive skills are better at strategically fulfilling their reading goals, and therefore should use highlighting more selectively (Burin et al., 2020). This finding is also inconsistent with a previous study suggesting that high-skilled readers highlight a significantly lower proportion of the text, and the text they highlight are significantly more relevant (Bell & Limber, 2010). Moreover, we found no relationship between the centrality of readers' highlighted text and comprehension performance. One possible explanation is that we measured metacognitive skills, an aspect of readers' skills that is different from what has been measured in the Bell and Limber (2010) study, which is vocabulary using the Nelson-Denny reading test. While both measures are often used to differentiate high-skilled and low-skilled readers, it is certainly possible that higher quality text-marking is associated with a higher score in vocabulary breadth, but not higher metacognitive skills. This inconsistency between our finding and previous literature may also be partly attributed to different characteristics of the text we used. The material used in the Bell and Limber (2010) study was an introductory psychology textbook, and the centrality of the texts were judged by course instructors who had clear standards of what students were expected to learn. The text we used are the ones more common in reading comprehension tests, with an average length of 458 words. These passages may have a higher information density compared to textbooks where lots of details, examples and illustrations can be incorporated. In other words, when readers highlight more words, specifically

more peripheral words, it might not be because they are not capable of distinguishing central and peripheral information, but because there is a higher proportion of important information in a 500-word passage.

Our second research question asked whether readers spend more time reading the text when they highlight. As predicted, readers who highlighted during the task spent more time reading the text, and follow-up analysis suggested that there is no significant relationship between reading time and the number of words readers highlighted. This result provided evidence that this increased reading time in the two highlighting conditions is not simply due to extra time spent on the physical act of highlighting. Rather, we speculate that it is the goal of highlighting and the decision-making process, where readers need to distinguish central information and peripheral information and determine what information to highlight, that increases their time spent on reading.

We also asked whether or not readers with higher metacognitive skills engage in more deliberate highlighting as opposed to mindless highlighting by looking at whether or not they slowed down more in the highlighting condition. Contrary to our hypothesis, there is no relationship between metacognitive skills and the time spent reading for readers in the two highlighting conditions. There are multiple potential reasons for this result. The most important factor is our choice of total reading time as a measure for the attention allocated to a part of the text. This research question is based on the assumption that deliberate highlighting requires more attention than mindless highlighting, hence requiring a longer reading time. However, readers with better metacognitive skills may also be more skillful readers and take less time to comprehend the text, which interferes with our use of reading time to measure attention. Future research should examine readers' behavior and reading speed before they highlight a part of a text.

The goal of our third research question was to examine whether or not text-marking was a beneficial strategy for reading comprehension, and whether its effect depends on the physical act of highlighting or the visual cues. We did not find any evidence that either the physical act of highlighting or the visual cues improved accuracy or question answering time. While the result is not what we expected, it echoes previous literature where text-marking had no effect on the accuracy of reading comprehension, and sometimes turned out to be counterproductive (Gier et al., 2009; Rickards & Denner, 1979; Peterson, 1991). Some literature ascribed this lack of significance to the low quality of reader-generated highlighting. However, we found this explanation to be an oversimplification, because as demonstrated in the follow-up analyses for the first question, we found no relationship between the centrality of readers' highlighting and their accuracy.

While we did not find evidence that text-marking reduced questions answering time, there is a trend that participants who had access to their own highlighted text spent a longer time on answering questions, compared to the two groups who answered questions with a clean text, although this trend did not reach significance. This might be due to the limited power of our study and the large variance in readers' question answering time. Unlike passage reading time, where we multiplied the average fixation time per word by the number of words in a passage to gain a rough measure of the minimal reading time, there is no valid measure to determine whether or not a participant was on-task while answering questions. Nonetheless, this trend pointed us to a future direction to examine whether or not the presence of markings in a text interferes with the speed of searching for information in a text, which is also explored through our final question.

Our final question focused on the group who had access to their highlighted text when answering questions, and asks whether or not more highlighting leads to a longer time spent on locating information. Contrary to our hypothesis, a larger proportion of highlighting and more

peripheral highlighting did not slow down readers' speed of searching for information. Meanwhile, this result should be viewed with caution because it remains unknown whether or not each reader rereads the text when answering questions. According to Yeari et al. (2017), the effects of highlighting only occurred in the re-reading stage and not during initial reading, meaning that the readers' attention allocation were not influenced by external cues when they processed a text for the first time. If readers did not go back to the text and instead answered questions based on what was stored in their working memory, then it was natural that the quantity and quality of text-marking did not interfere with question answering time.

This lack of effect is consistent with the result from a follow-analysis for our first question, where there was no significant relationship between the quality of text-marking and comprehension accuracy. This study is one of the few studies that used empirical methods to study the relationship between the quality of text-marking and comprehension performance, and the first to our best knowledge in the context of reader-generated text-markings. Together, these results drove us to reflect on the recommendations made in previous studies that readers should limit their text-marking under 20% or within one sentence per paragraph, and should prioritize marking conceptual sentences as opposed to subordinate ones. Marking central information is one of the ways students optimize their comprehension, but it may not be the only way. For example, List and Alexander (2020) investigated student's use of multiple texts, where students could use highlighting as an external support. When asked to explain their reasoning for highlighting a specific piece of information, some of the explanations given by students included that they were unfamiliar with the information, they found a piece of information difficult to understand, they were able to connect the information to a part of their prior knowledge, etc. Therefore, there are

lots of valid reasons why readers may highlight a piece of seemingly peripheral information while still engaging in metacognitive activities.

Limitations

One of the major limitations of this study is that without an eye-tracking methodology, we have little knowledge of the reading processes of participants when they use text-marking. This limited our understanding of our research questions on three dimensions: Firstly, and probably most importantly, we could not tell whether or not readers were giving their best effort completing the tasks. There was a large variance in the question answering time of our participants which reduced the power of our analyses. With an eye-tracking measure, we will be able to see how much of the passage and questions the readers read and examine whether or not this difference in engagement accounts for some of the variance in accuracy and question answering time. Secondly, we chose reading time as a measure to distinguish selective highlighting and mindless highlighting, but could not distinguish between readers who spent a longer time because they were less skillful at reading, and readers who spent more time at places where they intended to highlight in order to make a more deliberate decision. Eye-tracking measures allow us to examine if readers highlight information on their initial reads or re-reads, whether or not they reread previous sentences to determine the centrality of a new information, and how these behaviors are associated with their metacognitive skills. Finally, we could not tell whether or not readers reread the text to answer the questions, and question answering time is a very rough measure of readers' precision in locating relevant information in the text. With eye-tracking data we will be able to identify readers who reread the text when answering questions, and see how much time they spent on rereading the irrelevant section of the text before rereading the part where the answer is, which is a better measure of the precision of their rereads compared to the question answering time. We will then

be able to break down total question answering time and find how much time readers spend on answering each question, and whether or not they have better precision locating information for questions whose answers have already been highlighted in the text. In sum, total reading time and question answering time are very rough measures of readers' attention allocation and searching efficiency, both of which will be more accurately captured by eye-tracking measures.

Another important limitation concerns the validity and reliability of the metacognitive inventory we used. While self-reported metacognitive measures are cost-effective and easy to implement, their validity has been questioned because what students report can be their general perception about these strategies as opposed to how often and how well they use these strategies. Studies have also shown that readers have better accuracy at reporting concurrent measures, which are the strategies that they just used in their reading task, as opposed to prospective measures, the strategies they generally use, and they often overestimate the number of tactics they use (Cromley & Azevedo, 2006; Cromley & Azevedo, 2011; McCardle & Hadwin, 2015; Winne & Jamieson-Noel, 2002). Because many items on the metacognitive awareness inventory do not apply to the reading task in our study (e.g., using reference materials; discussing with peers about the material), we asked readers to report how often they use these strategies in reading in general, which could reduce the validity. Moreover, the reliability of each of the subscales are 0.574 (global reading strategies), 0.719 (problem-solving strategies) and 0.567 (supporting reading strategies), none of which is higher than 0.8. Overall, the readers' MARSI score may not capture their metacognitive ability accurately, and future studies should use alternatives that reflect metacognitive strategy uses more accurately, such as eye-tracking data, think-aloud protocols and error/inconsistency detection.

Future Directions

The effectiveness of text-marking on reading and learning is still a question under debate as past studies generated conflicted results. While our study adds to literature suggesting that text-marking did not yield significant benefits for reading comprehension, it is more likely that the effects of this strategy depends on many other factors, including text-length, text-difficulty, text-accessibility and criterion tasks. Therefore, in order to better interpret these conflicted results, future studies should investigate whether or not the effect of text-marking varies across different text-characteristics and learning conditions. Compared to a simple "yes/no" question on the effectiveness of text-marking, such questions will be more informative for educators and learners, given that reading in each discipline and under each context (e.g. reading a textbook to prepare for a test; reading a journal article for writing papers; reading an article on a comprehension test) have different expectations and require different strategies. For example, the majority of previous studies used cued recall and free recall as their criterion task, where reading and text-marking were used as ways of remembering information and preparing for an assessment (Fuson-Newsome & Metzger, 2016; Gier et al., 2009; Lorch et al., 1995; Yue et al., 2015). These designs also align with the von Restorff effect, indicating that text-marking would benefit reading and learning because the highlighted text looks different among other items, and therefore will be more readily recalled. In our study, however, the texts were accessible to readers during the assessment. There is less reliance on storing information in readers' working memory and recalling them, but more reliance on effective and efficient searching among the text. Therefore, a future study could provide half of the participants with a closed-book comprehension test and half of the participants with an open-book test, and examine whether or not text-marking benefits readers in the closed-book condition more than readers in the open-book condition.

Perhaps more important than the effectiveness of text-marking is how readers come to the decision of what to highlight and how they make use of these markings in rereading or learning. As we discussed earlier, even if text-marking is beneficial, it can become merely an external move without comprehension monitoring and regulation. Therefore, future studies could use either a think-aloud protocol or an eye-tracking measure to understand how readers decide which part of the text to mark, and whether or not they determine the relative importance of an information through rereading some of the texts or building connections between different parts of the texts. In addition, future studies could investigate whether or not the presence of text-marking, especially those relevant to questions on the test, increases the precision of readers' searching of information (i.e., Less time on the irrelevant areas before locating the answer).

Implications

Despite the popularity of text-marking as a learning strategy among readers, our study did not find evidence of its benefits, and can be more time consuming if we take the increased reading time into account. Meanwhile, just because there is not significant evidence of its benefits on a reading comprehension test does not mean that text-marking does not benefit other forms of learning and assessment. We recommend text-marking should be used with caution and educators should inform students of the empirical evidence on both sides, and teach students to make decisions on whether or not to use text-marking based on the nature of the text and task. Paris et al. (1983) discussed in their article "Becoming a Strategic Reader" that there are three types of knowledge around reading strategies: One of them is declarative knowledge, which is knowledge about the task structure, task requirement and readers' ability; The other type is conditional knowledge, which is knowledge about how to perform different reading strategies like skimming and summarizing. What is often overlooked, however, is what is called conditional knowledge,

which informs the reader of when those strategies should be used and why they are appropriate under those circumstances. Therefore, educators should aim to gain and teach conditional knowledge of text-marking and help students understand the mechanism behind why text-marking does or does not work.

Our study fills in some of the gaps in previous literature on the relationship between readers' characteristics and their strategy use. Specifically, we focused on readers' metacognitive skills and found that readers with higher metacognitive skills highlighted more peripheral words. Despite that the direction of our finding is the opposite of what we expected, this relationship showed that the use of text-marking differs across readers' characteristics. Therefore, the instruction of how to use text-marking more effectively would vary across each individual. This individualized adaptation is particularly important for college-level readers because of the complexity of readings in college. Simple instructions like "the first sentence of each paragraph is the topic sentence" does not align well with these texts and does not fulfill the learning demand of college students. In addition, we argued that some of the conventional recommendations on highlighting, such as highlighting the topic sentence and limiting highlighting to one sentence per paragraph, oversimplifies the roles highlighting can play in reading comprehension. Our study found no evidence that highlighting peripheral information had a detrimental effect on comprehension performance. A reader may highlight a piece of peripheral information because they found it hard to understand during the initial read, or because they built a connection between this information and a previous argument in the text or in their prior knowledge. These reasons all reflected high-level metacognitive activities, yet the product of which might be categorized as low-quality highlighting using conventions in past literature. Therefore, we recommend that instructors help students monitor their process of text-marking as opposed to its product. That is, instead of using

the centrality of their highlighted text themselves as a marker of how much metacognitive activities readers engage in, instructors should focus more on teaching students to use text-marking to better fulfill their reading goals.

Appendix A

METACOGNITIVE AWARENESS OF READING STRATEGIES INVENTORY-REVISED (MARSI-R, 2013)

Participant Number: _____

INSTRUCTIONS FOR COMPLETING THE INVENTORY

The statements listed on this inventory describes 15 strategies or actions readers use when reading

academic or school-related materials such as book chapters, journal articles, stories, etc.

Directions:

Step 1: Read each statement to indicate whether you are aware of and/or use these strategies

when you read.

Step 2: Use the following scale to show your strategy awareness and/or use:

1. I have never heard of this strategy before.

2. I have heard of this strategy, but I don't know what it means.

3. I have heard of this strategy, and I think I know what it means.

4. I know this strategy, and I can explain how and when to use it.

5. I know this strategy quite well, and I often use it when I read.

Step 3: After reading each strategy statement, place the numbers (1, 2, 3, 4, or 5) in the spaces

preceding each statement to show your level of awareness and/or use of each strategy.

Example:_____Sounding words out when reading

Place the number 1 in the blank space next to the strategy if you've never heard of it before; place

the number 2 next to the strategy if you've heard of it, but don't know what it means; and so on.

There are no right or wrong answers to the statements in this inventory. It takes about 7-10

minutes to complete the inventory.

METACOGNITIVE AWARENESS OF READING STRATEGIES INVENTORY-REVISED (MARSI-R, 2013)

Strategy scale:

1. I have never heard of this strategy before.
2. I have heard of this strategy, but I don't know what it means.
3. I have heard of this strategy, and I think I know what it means.
4. I know this strategy, and I can explain how and when to use it.
5. I know this strategy quite well, and I often use it when I read.

After reading each strategy statement, place the numbers (1, 2, 3, 4, or 5) in the spaces preceding each statement to show your level of awareness and/or use of each strategy.

Strategies 1-15

_____01. Having a purpose in mind when I read.

_____02. Taking notes while reading.

_____03. Previewing the text to see what it is about before reading it.

_____04. Reading aloud to help me understand what I'm reading.

_____05. Checking to see if the content of the text fits my purpose for reading.

_____06. Discussing what I read with others to check my understanding.

_____07. Getting back on track when getting sidetracked or distracted.

_____08. Underlining or circling important information in the text.

_____09. Adjusting my reading pace or speed based on what I'm reading.

_____10. Using reference materials such as dictionaries to support my reading.

_____11. Stopping from time to time to think about what I'm reading.

_____12. Using typographical aids like bold face and italics to pick out key information.

_____13. Critically analyzing and evaluating the information read.

_____14. Re-reading to make sure I understand what I'm reading.

_____15. Guessing the meaning of unknown words or phrases.

_____

GENERAL INFORMATION

Age: _____

Class Year: _____

Gender:_____

Race or Ethnicity:

_____African American/Black/ Afro-Caribbean

_____White/European

_____American Indian/Native American

_____Asian/Asian American

_____Hispanic/Latino/Latina/Latinx

_____Other (please specify)

I consider myself (Check one):

1. _____An excellent reader

2. _____A good reader

3. _____An average reader

4. _____A poor reader

Appendix B

Reading Comprehension Task

In this task, you will read a total of six passages. As you read, please use the highlighter function to highlight texts that you consider are central to the passage in yellow.

After each passage, there are five comprehension questions, and you can put down your choice on the line at the beginning of each question. You should not use additional highlighting while answering questions, or use any other functions apart from typing the answer.

There is no time limit for this task, but you will be asked to put down the time you start reading each passage, the time you finish reading the passage and the time you finish answering questions. The six passages should take around forty-five minutes to finish.

On the next passage, you will see a practice passage and five practice questions.

Participant Number: _____

Practice Trial

Please put down the time you begin reading:_____(hour):_____(minute):_____(second)

The past decade has seen a statistically significant uptick in reports of the bacterial strains known as "super-bugs," so called not because of enhanced virulence, but because of their resistance to many antimicrobial agents. In particular, researchers have become alarmed about NDM-1 (New Delhi metallo-beta-lactamase), which is not a single bacterial species, but a transmittable genetic element encoding multiple resistance genes. A resistance "cocktail" such as NDM-1 could bestow immunity to a bevy of preexisting drugs simultaneously, rendering the bacterium nearly impregnable.

However, in spite of the well-documented dangers posed by antibiotic-resistant bacteria, many scientists argue that the human race has more to fear from viruses. Whereas bacteria reproduce asexually through binary fission, viruses lack the necessary structures for reproduction, and so are known as "intracellular obligate parasites." Virus particles called virions must marshal the host cell's ribosomes, enzymes, and other cellular machinery in order to propagate. Once various viral components have been built, they bind together randomly in the cellular cytoplasm. The newly finished copies of the virus break through the cellular membrane, destroying the cell in the process. Because of this, viral infections cannot be treated ex post facto in the same way as bacterial infections, since antivirals designed to kill the virus could do critical damage to the host cell itself. In fact, viruses can infect bacteria (themselves complete cells), but not the other way around. For many viruses, such as that responsible for the common cold sore, remission rather than cure is the goal of currently available treatment.

While the insidious spread of drug-resistant bacteria fueled by overuse of antibiotics in agriculture is nothing to be sneezed at, bacteria lack the potential for cataclysm that viruses have. The prominent virologist Nathan Wolfe considers human immunodeficiency virus (HIV), which has resulted in the deaths of more than thirty million people and infected twice that number, "the biggest near-miss of our lifetime." Despite being the most lethal pandemic in history, HIV could have caused far worse effects. It is only fortunate happenstance that this virus cannot be transmitted through respiratory droplets, as can the pathogenic viruses that cause modern strains of swine flu (H1N1), avian flu (H5N1), and SARS.

Please put down the time you finish reading:_____(hour):_____(minute):_____(second)

_____1. The main purpose of the passage can be expressed most accurately by which of the following?

(A) To contrast the manner by which bacteria and viruses infect the human body and cause cellular damage

(B) To explain the operations by which viruses use cell machinery to propagate

(C) To argue for additional resources to combat drug-resistant bacteria and easily transmissible pathogenic viruses

(D) To highlight the good fortune experienced by the human race, in that the HIV pandemic has not been more lethal.

(E) To compare the relative dangers of two biological threats and judge one of them to be far more important.

_____2. It can be inferred from the passage that infections by bacteria

(A) result from asexual reproduction through binary fission

(B) can be treated ex post facto by antimicrobial agents, unlike viral infections

(C) can be rendered vulnerable by a resistance cocktail such as NDM-1

(D) are rarely cured by currently available treatments, but rather only put into remission

(E) mirror those by viruses, in that they can both do critical damage to the host cell

_____3. According to the passage, intracellular obligate parasites

(A) are unable to propagate themselves on their own

(B) assemble their components randomly out of virions

(C) reproduce themselves through sexual combination with host cells

(D) have become resistant to antibiotics through the overuse of these drugs

(E) construct necessary reproductive structures out of destroyed host cells

_____4. The author believed that the spread of antibiotic-resistant bacteria is caused by:

(A) The emergence of NDM-1

(B) Overuse of antibiotics in agriculture

(C) Lack of accessibility of other drugs

(D) Lack of documentation of its danger

_____5. According to the author, what is a major difference between the HIV and SARS?

(A) There is no drug available for treating SARS

(B) SARS can be transmitted through respiratory droplets

(C) SARS infected more people compared to HIV

(D) SARS was given less attention by virologists

Please put down the time you finish answering questions:
____(hour):____(minute):____(second)

You cannot change your answers once you move on to the next passage.

Passage 1

Please put down the time you begin reading:_____(hour):_____(minute):___(second)

Most recent work on the history of leisure in Europe has been based on the central hypothesis of a fundamental discontinuity between preindustrial and industrial societies. According to this view, the modern idea of leisure did not exist in medieval and early modern Europe: the modern distinction between the categories of work and leisure was a product of industrial capitalism. Preindustrial societies had festivals (together with informal and irregular breaks from work), while industrial societies have leisure in the form of weekends and vacations. The emergence of leisure is therefore part of the process of modernization. If this theory is correct, there is what Michel Foucault called a conceptual rupture between the two periods, and so the very idea of a history of leisure before the Industrial Revolution is an anachronism.

To reject the idea that leisure has had a continuous history from the Middle Ages to the present is not to deny that late medieval and early modern Europeans engaged in many pursuits that are now commonly considered leisure or sporting activities— jousting, hunting, tennis, card playing, travel, and so on—or that Europe in this period was dominated by a privileged class that engaged in these pursuits. What is involved in the discontinuity hypothesis is the recognition that the people of the Middle Ages and early modern Europe did not regard as belonging to a common category activities (hunting and gambling, for example) that are usually classified together today under the heading of leisure. Consider fencing: today it may be considered a "sport," but for the gentleman of the Renaissance it was an art or science. Conversely, activities that today may be considered serious, notably warfare, were often described as pastimes.

Serious pitfalls, therefore, confront historians of leisure who assume continuity and who work with the modern concepts of leisure and sport, projecting them back onto the past without asking about the meanings contemporaries gave to their activities. However, the discontinuity hypothesis can pose problems of its own. Historians holding this view attempt to avoid anachronism by means of a simple dichotomy, cutting European history into two eras, preindustrial and industrial, setting up the binary opposition between a "festival culture" and a "leisure culture." The dichotomy remains of use insofar as it reminds us that the rise of industrial capitalism was not purely a phenomenon of economic history, but had social and cultural preconditions and consequences. The dichotomy, however, leads to distortions when it reduces a great variety of medieval and early modern European ideas, assumptions, and practices to the simple formula implied by the phrase "festival culture."

Please put down the time you finish reading:_____(hour):____(minute):___(second)

_____1. The primary purpose of the passage is to

(A) refute the idea that the history of leisure is discontinuous

(B) show why one of two approaches is more useful in studying the history of leisure

(C) suggest the need for a new, more inclusive concept to replace the concept of leisure

(D) trace the development of a theory about the history of leisure

(E) point out the basis for, and the limits of, an approach to the history of leisure

_____2. The author of the passage asserts that the "dichotomy" (line 24) can lead to which of the following?

(A) Reliance on only one of several equally valid theoretical approaches

(B) The imposition of modern conceptions and meanings on past societies

(C) Failure to take into account the complexity of certain features of European culture

(D) Failure to utilize new conceptual categories in the study of the history of leisure

(E) Failure to take account of the distinction between preindustrial and industrial societies

_____3. According to the passage, the "simple dichotomy" (line 24) is useful primarily because it serves as

(A) a way of calling historians' attention to certain facts about the Industrial Revolution

(B) an antidote to the oversimplification encouraged by such terms as "festival culture"

(C) a device for distinguishing between the work and the leisure activities of preindustrial Europeans

(D) a way of understanding the privileged class of medieval Europe by viewing its activities in modern terms

(E) a tool for separating social history, including the history of leisure, from economic history

_____4. Which of the following best describes the organization of the passage as a whole?

(A) Two hypotheses are discussed, and evidence in support of one is presented.

(B) A hypothesis is presented and discussed, and a limitation to the hypothesis is identified.

(C) A hypothesis is proposed, its supposed advantages are shown to be real, and its supposed disadvantages are shown to be illusory.

(D) A problem is identified, two hypotheses are advanced to resolve it, and both are rejected.

(E) A problem is identified, two resolutions are proposed, and a solution combining elements of both is recommended.

_____5. What is an example from the passage that was considered leisure in the past but not now?

(A) Science

(B) Art

(C) Hunting

(D) Fencing

(E) Warfare

Please put down the time you finish answering questions:
____(hour):____(minute):____(second)

You cannot change your answers once you move on to the next passage.

Passage 2

Please put down the time you begin reading:_____(hour):_____(minute):___(second)

The story of Sir Gawain and the Green Knight has its foundation in Arthurian legend as formulated and passed down by the pagan oral tradition. In its written form, however, the tale bears the marks of Christian influence—it contains numerous scriptural and doctrinal references to Christianity. Since the author of Sir Gawain and the Green Knight is unknown, it is difficult to determine with any certainty the extent to which he was responsible for the incorporation of Christianity into the legend. For all we know, the story may have been "Christianized" in its oral form long before the poet set it into writing. The poet himself supports this possibility by writing in the opening lines that he will tell "anew" the tale "as I heard it in hall." If this is the case (and even if it is not), it is distinctly possible that the heroes of the Arthurian tradition represent in the written form a pagan interpretation of Christian ideals, rather than an externally imposed Christianization of pagan codes of behavior.

While it could certainly be argued that the poet portrays Sir Gawain as a good Christian hero in an attempt to infuse the story with Christian values, the critical tone of the narrative seems to suggest a different conclusion—that by critically editorializing the paganized form of

Christianity embodied by Sir Gawain, the poet is trying to correct what he sees to be the flaws of that form. From the perspective of this conclusion it is clear that the poet only "Christianizes" the traditional legend to the extent that he criticizes the pagan interpretation of Christianity that is inherent in the behavior of its heroes.

Those who would argue that the poet intends to portray Sir Gawain as the perfect Christian hero would point to the descriptions of his chivalric qualities. The poet does indeed describe Gawain's Christian virtues generously; he even makes a special aside early in the second fit to describe the significance of the pentangle embossed on Gawain's shield, and to explain "why the pentangle is proper to that peerless prince." The author then delves into a lengthy enumeration of Gawain's Christian virtues. What is more, the fact that he uses the pentangle—a pagan symbol—to do it would seem to suggest that the author does indeed intend to add a Christian interpretation to the pagan legend he is retelling. Taken in its larger context, however, this passage takes on a different significance. In further examination of the poet's descriptions of Sir Gawain, it becomes apparent that the knight's seemingly perfect Christian behavior is superficial. A contrast can be observed between his "Christian" words and actions and his decidedly un-Christian motives. One theory is that, by emphasizing this contrast, the poet intends to denounce the pagan "misunderstanding" of the Christian message.

Please put down the time you finish reading:____(hour):____(minute):____(second)

_____1. Which of the following can be inferred about the pagan and Christian origins of Sir Gawain and the Green Knight?

(A) As an orally-handed-down tale, it was pagan, but as a written tale, it was Christian.

(B) Sir Gawain was a knight in King Arthur's court.

(C) Sir Gawain and the Green Knight contains both Christian and pagan elements, although it is not clear that either perspective is dominant.

(D) Both A and B

(E) A, B and C

_____2. Which of the following can be inferred from the author's interpretation of the Christian aspects of the poem presented in the third paragraph?

(A) Pagans and Christians differ in their interpretations of the Christian symbolism in the story.

(B) A pagan cannot have motives that are acceptable from a Christian perspective.

(C) A pagan story cannot be used to convey a Christian attitude.

(D) Christianity was absent in Arthurian stories before such stories were written down.

(E) Being a good Christian involves having both the right actions and the right motives.

_____3. Which of the following, if true, would most undermine the "theory" mentioned in the final sentence of the passage?

(A) Sir Gawain is portrayed as disingenuous in his exercise of "Christian virtues."

(B) Another character in the story is also associated with pagan symbols and is praised straightforwardly for her Christian virtues.

(C) Sir Gawain, in the story, prays to God to help him in battle.

(D) Another character in the story is associated with pagan symbols but is portrayed as having no Christian virtues whatsoever.

(E) A group of people in the story are portrayed as "barbarians" who are neither pagan nor Christian.

_____4. What do we know about the author of Sir Gawain and the Green?

(A) The story has multiple authors

(B) The author is unknown

(C) The author used a pseudonym when writing the story

(D) The author was responsible for including Christianity in the story

_____5. Where is the pentangle that the poet depicted?

(A) On Gawain's sword

(B) On Gawain's shield

(C) On Gawain's armour

(D) On Gawain's helmet

Please put down the time you finish answering questions:
____(hour):____(minute):____(second)

You cannot change your answers once you move on to the next passage.

Passage 3

Please put down the time you begin reading:_____(hour):_____(minute):___(second)

In traditional theater forms, the roles of performer and audience are completely separate, so that performance space can be said to encompass an actors' sphere and a spectators' sphere. Even when performers move out into the audience or when there is scripted audience interaction, spectators do not become performers. Finally, while stories may open up the imagination or excite audiences, according to Augusto Boal, they discourage political action by providing catharsis. The passive spectator follows the play's emotional arc and, once the action concludes, finds the issue closed. Boal reminds us that our theater etiquette creates a kind of culture of apathy where individuals do not act communally, despite shared space, and remain distanced from art.

Workshop theater, such as Boal's Image Theatre and Forum Theatre, is a response to that. In the workshop form, performance space is created for a select group of people, but the performers' sphere and the audience's sphere are collapsed: everyone is at once theater maker and witness. In Image Theatre, participants will come up with a theme or issue and arrange themselves into a tableau that depicts what that issue looks like in society today, versus what the ideal situation would be. They then try to transition from the current image to the ideal image in a way that seems plausible to all the participants. Forum Theatre, on the other hand, creates a narrative skit depicting a certain problem. After the actors have gone through the action of the play once, a facilitator, known as the joker (like the one in a pack of cards), encourages those who have watched the story to watch it again and to stop it at any time to take the place of the protagonist. The aim is to find a solution to the problem, realizing along the way all of the obstacles involved. In Forum Theatre, just as in Image Theatre, there is not always a solution. The main goal of this form, then, is to engage in the action, to reflect, and to understand particular issues as being part of a larger picture, thus using art to re-cast what seem like private troubles in a public, political light.

The main reason Boal developed these workshop styles was to grant audiences agency so that they may create ways to free themselves of oppression. Because he found theater audiences to be locked into a passive role—just like he found the oppressed coerced into a subservient role in relation to their oppressors—he created the "spectators," or someone who simultaneously witnesses and creates theater.

Please put down the time you finish reading:____(hour):____(minute):____(second)

_____1.The second paragraph of the passage serves to

(A) elaborate on the topic of the first paragraph

(B) provide a rationale for an artistic endeavor

(C) discuss an artistic answer to a passive culture

(D) explain the theater's lack of appeal

(E) evaluate two contrasting styles of theater

_____2.The author uses the word agency to mean

(A) profit

(B) organization

(C) publicity

(D) power

(E) hegemony

_____3.Which of the following would Boal consider a "spectactor"?

(A) a person who engages in political action

(B) an audience member who finds catharsis in a play

(C) any person placed in a subservient role

(D) any actor

(E) a participant in an Image workshop

_____4.According to Boal, all of the following are disadvantage of traditional theater forms EXCEPT:

(A) Such productions prevent the actors from going into the audience.

(B) Such productions provide catharsis.

(C) Such productions discourage communal activity.

(D) Such productions obstruct political change.

(E) Such productions distance the audience from the art.

_____5.All of the following would be characteristic of a Forum workshop EXCEPT:

(A) Productions begin with a narrative script.

(B) Different people often play the protagonist.

(C) Some performances do not achieve catharsis.

(D) Participants arrange themselves into a tableau.

(E) Performances are guided by a mediator.

Please put down the time you finish answering questions:
____(hour):____(minute):____(second)

You cannot change your answers once you move on to the next passage.

Passage 4

Please put down the time you begin reading:_____(hour):____(minute):___(second)

Until recently, many anthropologists assumed that the environment of what is now the southwestern United States shaped the social history and culture of the region's indigenous peoples. Building on this assumption, archaeologists asserted that adverse environmental conditions and droughts were responsible for the disappearances and migrations of southwestern populations from many sites they once inhabited.

However, such deterministic arguments fail to acknowledge that local environmental variability in the Southwest makes generalizing about that environment difficult. To examine the relationship between environmental variation and sociocultural change in the Western Pueblo region of central Arizona, which indigenous tribes have occupied continuously for at least 800 years, a research team recently reconstructed the climatic, vegetational, and erosional cycles of past centuries. The researchers found it impossible to provide a single, generally applicable characterization of environmental conditions for the region. Rather, they found that local areas experienced different patterns of rainfall, wind, and erosion, and that such conditions had prevailed in the Southwest for the last 1,400 years. Rainfall, for example, varied within and between local valley systems, so that even adjacent agricultural fields can produce significantly different yields.

The researchers characterized episodes of variation in southwestern environments by frequency: low-frequency environmental processes occur in cycles longer than one human generation, which generally is considered to last about 25 years, and high frequency processes have shorter cycles. The researchers pointed out that low-frequency processes, such as fluctuations in stream flow and groundwater levels, would not usually be apparent to human populations. In contrast, high-frequency fluctuations such as seasonal temperature variations are observable and somewhat predictable, so that groups could have adapted their behaviors accordingly. When the researchers compared sequences of sociocultural change in the Western Pueblo region with episodes of low- and high-frequency environmental variation, however, they found no simple correlation between environmental process and sociocultural change or persistence.

Although early Pueblo peoples did protect themselves against environmental risk and uncertainty, they responded variously on different occasions to similar patterns of high-frequency climatic and environmental change. The researchers identified seven major adaptive responses, including increased mobility, relocation of permanent settlements, changes in subsistence foods, and reliance on trade with other groups. These findings suggest that groups' adaptive choices depended on cultural and social as well as environmental factors and were flexible strategies rather than uncomplicated reactions to environmental change. Environmental conditions mattered, but they were rarely, if ever, sufficient to account for sociocultural persistence and change. Group size and composition, culture, contact with other groups, and

individual choices and actions were — barring catastrophes such as floods or earthquakes — more significant for a population's survival than were climate and environment.

Please put down the time you finish reading:____(hour):____(minute):____(second)

_____1. The passage is primarily concerned with

(A) explaining why certain research findings have created controversy

(B) pointing out the flaws in a research methodology and suggesting a different approach

(C) presenting evidence to challenge an explanation and offering an alternative explanation

(D) elucidating the means by which certain groups have adapted to their environment

(E) defending a long-held interpretation by presenting new research findings

_____2. Which of the following findings would most strongly support the assertion made by the archaeologists mentioned in line 3?

(A) A population remained in a certain region at least a century after erosion wore away much of the topsoil that sustained grass for their grazing animals.

(B) The range of a certain group's agricultural activity increased over a century of gradual decrease in annual rainfall.

(C) As winters grew increasingly mild in a certain region, the nomadic residents of the region continued to move between their summer and winter encampments.

(D) An agricultural population began to trade for supplies of a grain instead of producing the grain in its own fields as it had in the past.

(E) A half century of drought and falling groundwater levels caused a certain population to abandon their settlements along a riverbank.

_____3. The fact that "adjacent agricultural fields can produce significantly different yields" (lines 16–17) is offered as evidence of the

(A) unpredictability of the climate and environment of the southwestern United States

(B) difficulty of producing a consistent food supply for a large population in the Western Pueblo region

(C) lack of water and land suitable for cultivation in central Arizona

(D) local climatic variation in the environment of the southwestern United States

(E) high-frequency environmental processes at work in the southwestern United States

_____4. It can be inferred from the passage that which of the following activities is NOT an example of a population responding to high-frequency environmental processes?

(A) Developing watertight jars in which to collect and store water during the rainy season

(B) Building multistory dwellings in low-lying areas to avoid the flash flooding that occurs each summer

(C) Moving a village because groundwater levels have changed over the last generation

(D) Trading with other groups for furs from which to make winter clothes

(E) Moving one's herds of grazing animals each year between summer and winter pastures

_____5. How are low-frequency climate changes differentiated from high-frequency climate changes?

(A) High-frequency climate change occurs every year

(B) High-frequency climate change has cycles shorter than a human generation

(C) High-frequency climate change is not as noticeable for human

(D) High-frequency climate change predicts a significant portion of socio-cultural changes

Please put down the time you finish answering questions:
____(hour):____(minute):____(second)

You cannot change your answers once you move on to the next passage.

Passage 5

Please put down the time you begin reading:_____(hour):_____(minute):_____(second)

Matisse and Picasso; Picasso and Matisse. Throughout the twentieth century, this pairing has been touted as the quintessential artistic rivalry. In *Matisse* and *Picasso*, Yve-Alain Bois follows Hubert Damisch in proposing that the interaction between Picasso and Matisse should be seen as

a dynamic game rather than a static conflict of artistic polarities. Bois employs the metaphor of chess, arguing that the game represents the artists' exchange as "a competitive rivalry and a complex temporality" that can be viewed both as a linear process and a simultaneous structure.

But the metaphor of a competitive sport, however complex and intellectually rich, is misleading. The two artists were engaged not just in competition (even friendly competition) but also in friendly dialogue. The two men were more than rivals: they were colleagues, critics, teachers, and occasional friends. A better model, though perhaps one with less flash, is that of a simple conversation, with all the rich variation and shifts in motivation and tone that are possible.

Picasso's *Large Nude* in a *Red Armchair* marks the extremes of the artist's combativeness towards Matisse. The painting is a clear parody of Matisse's earlier Odalisque with a Tambourine. The composition of the figures is strikingly similar: a woman lounges in an armchair at the center of the painting, arm raised above her head, decorative wallpaper behind her. Both paintings feature vivid color contrasts, with green wallpaper, vivid reds, glaring yellows, and rich browns. But Picasso's painting, finished in 1929, mocks the achievements of Matisse's earlier work. The sensuous, rich mood of Matisse's painting has been transformed in Picasso's work into something harsh and grotesque.

The other extreme of the dialogue between the two artists can be seen in Picasso's Woman with Yellow Hair and Matisse's response, The Dream. The exchange begins with Picasso's work, in 1931. The painting depicts a woman asleep on her arms, resting on a table. She is full, rich, warm, and curved, her head and arms forming a graceful arabesque. This image seems a direct attempt to master Matisse's style and to suggest to the older artist new directions for his work. While there may well be an edge of competitiveness to the painting, a sense that Picasso was demonstrating his ability to do Matisse's work, it remains in large part a helpful hint.

Matisse, nearly a decade later, continues the conversation in a similar tone. In The Dream of 1940, he proposes a revision of Picasso's work. Again, a woman lies asleep on a table, her arm tucked beneath her head. Matisse accepts Picasso's basic suggestions for his style: sinuous curves, volumes, and shocking uses of color to express an effect. But Matisse also modifies the earlier work significantly. Color is no longer rigidly tied to form, as bits of fuchsia seep outside the thick black line marking the outline of the table and the patch of yellow on the woman's blouse refuses to be contained by the drawn line. Matisse uses Picasso's same palette of red, purple, white, black, and yellow to create this revision, editing out only the garish green, as if to chide Picasso for the choice. The brilliant interplay of colors in Matisse's work is far more sophisticated and subtle than that offered by Picasso. "Thank you," Matisse seems to be saying, "but you missed a few spots."

Please put down the time you finish reading:____(hour):____(minute):____(second)

_____1. The primary purpose of the passage is to

(A) discuss the two best painters of an epoch

(B) evaluate a theory and endorse a revision

(C) compare selected works of two masters

(D) show that Matisse's work is more sophisticated

(E) illustrate how Picasso taught Matisse

_____2. The author would most likely agree with which of the following statements?

(A) Artistic rivalries are more like Olympic competitions than professional sports.

(B) Artistic mastery is best demonstrated by employing multiple styles.

(C) Artists must be good conversationalists.

(D) Artistic rivalries can actually be reciprocally nourishing.

(E) Artistic rivalries generally last for decades.

_____3. According to the passage, which of the following describes Woman with Yellow Hair?

(A) It was parody of a work by Matisse.

(B) Its colors were not rigidly tied to its form.

(C) Its color palette was larger than that of The Dream.

(D) It was a response to a work by Matisse.

(E) It was harsh and grotesque.

_____4. Which of the following, had it actually occurred during the artists' lifetimes, would further support the author's thesis?

(A) A joint exhibition of the two artists' work

(B) A radio broadcast of the two artists discussing painting

(C) A movie that dramatized the competition between the two artists

(D) A play that depicted the two artists playing chess

(E) A painting of the two artists

_____5. Matisse may paint Odalisque with a Tambourine in

(A) 1925

(B) 1930

(C) 1931

(D) 1935

(E) 1940

Please put down the time you finish answering questions:
____(hour):____(minute):____(second)

You cannot change your answers once you move on to the next passage.

Passage 6

Please put down the time you begin reading:_____(hour):_____(minute):_____(second)

Walls and wall building have played a very important role in Chinese culture. These people, from the dim mists of prehistory have been wall-conscious; from the Neolithic period – when ramparts of pounded earth were used - to the Communist Revolution, walls were an essential part of any village. Not only towns and villages; the houses and the temples within them were somehow walled, and the houses also had no windows overlooking the street, thus giving the feeling of wandering around a huge maze. The name for "city" in Chinese (ch'eng) means wall, and over these walled cities, villages, houses and temples presides the god of walls and mounts, whose duties were, and still are, to protect and be responsible for the welfare of the inhabitants. Thus a great and extremely laborious task such as constructing a wall, which was supposed to run throughout the country, must not have seemed such an absurdity.

However, it is indeed a common mistake to perceive the Great Wall as a single architectural structure, and it would also be erroneous to assume that it was built during a single dynasty. For the building of the wall spanned the various dynasties, and each of these dynasties somehow contributed to the refurbishing and the construction of a wall, whose foundations had been laid many centuries ago. It was during the fourth and third century B.C. that each warring state started building walls to protect their kingdoms, both against one another and against the northern nomads. Especially three of these states: the Ch'in, the Chao and the Yen, corresponding respectively to the modern provinces of Shensi, Shanzi and Hopei, over and above building walls that surrounded their kingdoms, also laid the foundations on which Ch'in Shih Huang Di would build his first continuous Great Wall.

The role that the Great Wall played in the growth of Chinese economy was an important one. Throughout the centuries many settlements were established along the new border. The garrison troops were instructed to reclaim wasteland and to plant crops on it, roads and canals were built, to mention just a few of the works carried out. All these undertakings greatly helped to increase the country's trade and cultural exchanges with many remote areas and also with the southern, central and western parts of Asia – the formation of the Silk Route. Builders, garrisons, artisans, farmers and peasants left behind a trail of objects, including inscribed tablets, household articles, and written work, which have become extremely valuable archaeological evidence to the study of defense institutions of the Great Wall and the everyday life of these people who lived and died along the wall.

Please put down the time you finish reading:____(hour):____(minute):____(second)

_____1. Chinese cities resembled a maze

(A) because they were walled.

(B) because the houses has no external windows.

(C) because the name for cities means 'wall'.

(D) because walls have always been important there.

_____2. Constructing a wall that ran the length of the country

(A) honoured the god of walls and mounts.

(B) was an absurdly laborious task.

(C) may have made sense within Chinese culture.

(D) made the country look like a huge maze.

_____3. The Great Wall of China

(A) was built in a single dynasty.

(B) was refurbished in the fourth and third centuries BC.

(C) used existing foundations.

(D) was built by the Ch'in, the Chao and the Yen.

_____4. Crops were planted

(A) on wasteland.

(B) to reclaim wasteland.

(C) on reclaimed wasteland.

(D) along the canals.

_____5. The Great Wall

(A) helped build trade only inside China.

(B) helped build trade in China and abroad.

(C) helped build trade only abroad.

(D) helped build trade only to remote areas.

Please put down the time you finish answering questions:
____(hour):____(minute):____(second)

You cannot change your answers once you move on to the next passage.

You have reached the end of this task. Thank you for your participation!

Appendix C

Example Information Unit Breakdown

Passage	Paragraph	Information Unit
Leisure	1	Most recent work on the history of leisure in Europe has been based on the central hypothesis of a fundamental discontinuity between preindustrial and industrial societies.
Leisure	1	According to this view, the modern idea of leisure did not exist in medieval and early modern Europe:
Leisure	1	the modern distinction between the categories of work and leisure was a product of industrial capitalism.
Leisure	1	Preindustrial societies had festivals (together with informal and irregular breaks from work),
Leisure	1	while industrial societies have leisure in the form of weekends and vacations.
Leisure	1	The emergence of leisure is therefore part of the process of modernization.
Leisure	1	If this theory is correct,
Leisure	1	there is what Michel Foucault called a conceptual rupture between the two periods,
Leisure	1	and so the very idea of a history of leisure before the Industrial Revolution is an anachronism.
Leisure	2	To reject the idea that leisure has had a continuous history from the Middle Ages to the present is not to deny
Leisure	2	that late medieval and early modern Europeans engaged in many pursuits that are now commonly considered leisure or sporting activities—
Leisure	2	jousting, hunting, tennis, card playing, travel, and so on—
Leisure	2	or that Europe in this period was dominated by a privileged class that engaged in these pursuits.
Leisure	2	What is involved in the discontinuity hypothesis is the recognition
Leisure	2	that the people of the Middle Ages and early modern Europe did not regard as belonging to a common category activities
Leisure	2	(hunting and gambling, for example)

Leisure	2	that are usually classified together today under the heading of leisure.
Leisure	2	Consider fencing:
Leisure	2	today it may be considered a “sport,”
Leisure	2	but for the gentleman of the Renaissance it was an art or science.
Leisure	2	Conversely, activities that today may be considered serious, notably warfare, were often described as pastimes.
Leisure	3	Serious pitfalls, therefore, confront historians of leisure who assume continuity and who work with the modern concepts of leisure and sport,
Leisure	3	projecting them back onto the past without asking about the meanings contemporaries gave to their activities.
Leisure	3	However, the discontinuity hypothesis can pose problems of its own.
Leisure	3	Historians holding this view attempt to avoid anachronism by means of a simple dichotomy,
Leisure	3	cutting European history into two eras, preindustrial and industrial,
Leisure	3	setting up the binary opposition between a “festival culture” and a “leisure culture.”
Leisure	3	The dichotomy remains of use insofar as it reminds us that the rise of industrial capitalism was not purely a phenomenon of economic history,
Leisure	3	but had social and cultural preconditions and consequences.
Leisure	3	The dichotomy, however, leads to distortions
Leisure	3	when it reduces a great variety of medieval and early modern European ideas, assumptions, and practices to the simple formula implied by the phrase “festival culture.”

Appendix D

MOUNT HOLYOKE COLLEGE INFORMED CONSENT FORM

Study: College Students' Reading Comprehension

Investigators: Jiaan Shang faculty advisor: Professor Katherine Binder

This project has been approved by the Institutional Review Board of Mount Holyoke College. The following informed consent is required by Mount Holyoke College for all participants in human subjects research.

Procedures: The procedures to be followed in the project will be explained to you, and any questions you may have about the aims or methods of the project will be answered.

In this study, we will give you a document with six passages. You will be asked to read the passages and answer five multiple-choice questions after each passage. There is no time limit for this task, but you will be asked to record the time you begin reading, the time you finish reading and the time you finish answering questions for each passage. Then, you will take a survey that asks you about your use of reading strategies. The entire study should take about an hour.

Risks of Being in the Study:

This is a minimal risk study. However, if you become uncomfortable at any point, you can stop participating in the study or skip a question without penalty.

Benefits of Being in the Study:

If you sign up for this study through the SONA system, you can receive SONA credits for an eligible psychology course you choose. If not, you can participate in a lottery and have the chance to win a gift card that is worth $10.

Confidentiality:

All data in this study will be treated as strictly confidential. Your name will never be identified with any of the information you provide; instead, your information will be identified with a number. Only the principal investigator of the project will have access to the list of participants' names and identification numbers.

The results of this study will be made part of a senior thesis and may be used in papers submitted for publication or presented at professional conferences, but under no circumstances will your name or other identifying characteristics be included.

If you provide your address in order to receive a report of this research upon its completion, that information will not be used to identify you in the data.

Research records will be kept on a Google drive that only the principal investigator, faculty advisor and lab members who are a part of this project have access to.

Voluntary Nature of the Study:

Your participation is voluntary. If you decide to participate, you are free to end your participation at any time without penalty. You may also choose to not answer specific questions and you may choose to not participate in specific parts of the study.

Contacts and Questions:

If you have any questions about the study, you can contact Jiaan Shang through email(shang22j@mtholyoke.edu) and phone(413-362-5216).

In addition, you may contact the Mount Holyoke College Institutional Review Board (institutional-review-board@mtholyoke.edu) for information about the rights of human subjects at Mount Holyoke-approved research.

Statement of Consent:

I have read and understand the above information. I am 18 years of age or older. I consent to participate in the study. I allow my data to be used for research purposes.

Participant Signature_____

Printed Name_____

Date_____

Would you like a report of the group results of this research project upon its completion?

Yes No

If yes, please provide an email or physical address to which the report should be sent:

_____

References

Afflerbach, P., & Cho, B. Y. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. *Handbook of research on reading comprehension*, 69-90.

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The reading teacher*, *61*(5), 364-373.

Bell, K. E., & Limber, J. E. (2010). Reading skill, textbook marking, and course performance. *Literacy Research and Instruction, 49*(1), 56–67. doi:10.1080/19388070802695879

Berry, T., Cook, L., Hill, N., & Stevens, K. (2011). An Exploratory Analysis of Textbook Usage and Study Habits: Misperceptions and Barriers to Success. *College Teaching*, 59(1), 31–39. http://www.jstor.org/stable/41305102

Boulware-Gooden, R., Carreker, S., Thornhill, A., & Joshi, R. M. (2007). Instruction of metacognitive strategies enhances reading comprehension and vocabulary achievement of third-grade students. *The reading teacher*, *61*(1), 70-77.

Broughton, S. H., Sinatra, G. M., & Reynolds, R. E. (2010). The nature of the refutation text effect: An investigation of attention allocation. *The Journal of Educational Research*, 103(6), 407–423. doi:10.1080/00220670903383101

Burin, D. I., Gonzalez, F. M., Barreyro, J. P., & Injoque-Ricle, I. (2020). Metacognitive regulation contributes to digital text comprehension in E-learning. *Metacognition and Learning*, *15*, 391-410. doi:10.1007/s11409-020-09226-8

Clump, M. A., Bauer, H., & Bradley, C. (2004). The Extent to which Psychology Students Read Textbooks: A Multiple Class Analysis of Reading across the Psychology Curriculum. *Journal of Instructional Psychology*, 31(3).

Connor, C. M., Radach, R., Vorstius, C., Day, S. L., McLean, L., & Morrison, F. J. (2015). Individual differences in fifth graders' literacy and academic language predict comprehension monitoring development: An eye-movement study. *Scientific Studies of Reading*, 19(2), 114-134.

Cromley, J. G., & Azevedo, R. (2006). Self-report of reading comprehension strategies: What are we measuring?. *Metacognition and Learning*, 1(3), 229-247.

Cromley, J., & Azevedo, R. (2011). Measuring strategy use in context with multiple-choice items. *Metacognition and Learning*, 6(2), 155-177.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4-58. doi:10.1177/1529100612453266

Fesel, S. S., Segers, E., de Leeuw, L., & Verhoeven, L. (2016). Strategy training and mind-mapping facilitates children's hypertext comprehension. *Written Language & Literacy*, 19(2), 131-156.

Gier, V. S., Herring, D., Hudnell, J., Montoya, J., & Kreiner, D. S. (2010). Active reading procedures for moderating the effects of poor highlighting. *Reading Psychology*, 31(1), 69-81.

Gier, V. S., Kreiner, D. S., & Natz-Gonzalez, A. (2009). Harmful effects of preexisting inappropriate highlighting on reading comprehension and metacognitive accuracy. *The Journal of general psychology*, *136*(3), 287-302. doi:10.3200/GENP.136.3.287-302

Goldman, S. R., Snow, C., & Vaughn, S. (2016). Common themes in teaching reading for understanding: Lessons from three projects. *Journal of Adolescent & Adult Literacy*, 60(3), 255-264.

Gurung, R. A., Weidert, J., & Jeske, A. (2010). Focusing on how students study. *Journal of the Scholarship of Teaching and Learning*, *10*(1), 28-35.

Hong-Nam, K., Leavell, A. G., & Maher, S. (2014). The relationships among reported strategy use, metacognitive awareness, and reading achievement of high school students. *Reading Psychology*, 35(8), 762-790.

List, A., & Alexander, P. A. (2020). Strategy use in learning from multiple texts: An investigation of the integrative framework of learning from multiple texts. In *Frontiers in Education* (Vol. 5, p. 205). Frontiers.

Lorch, E. P., & Klusewitz, M. A. (1995). Effects of typographical cues on reading and recall of text. *Contemporary educational psychology*, *20*(1), 51-64. doi:10.1006/ceps.1995.1003

McCardle, L., & Hadwin, A. F. (2015). Using multiple, contextualized data sources to measure learners' perceptions of their self-regulated learning. *Metacognition and Learning*, 10(1), 43-75.

McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational psychology review*, 19(2), 113-139.

Metzger, S., & Fuson-Newsome, S. (2016). Highlighting for reading comprehension: Is two better than one?. *The Journal of Undergraduate Research at Ohio State*, *6*.

Miller, B. W. (2015). Using reading times and eye-movements to measure cognitive engagement. *Educational Psychologist*, 50(1), 31–42. doi: 10.1080/00461520.2015.1004068

Miyatsu, T., Nguyen, K., & McDaniel, M. A. (2018). Five popular study strategies: Their pitfalls and optimal implementations. *Perspectives on Psychological Science*, *13*(3), 390-407.

Mokhtari, K., Dimitrov, D. M., & Reichard, C. A. (2018). Revising the metacognitive awareness of reading strategies inventory (marsi) and testing for factorial invariance. *Studies in Second Language Learning and Teaching*, *8*(2), 219-246. doi: 10.14746/ssllt.2018.8.2.3

Nist, S. L., & Hogrebe, M. C. (1987). The role of underlining and annotating in remembering textual information. *Literacy Research and Instruction*, *27*(1), 12-25. doi: 10.1080/19388078709557922

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary educational psychology*, 8(3), 293-316.

Peterson, S. E. (1991). The cognitive functions of underlining as a study technique. *Literacy Research and Instruction*, *31*(2), 49-56. doi: 10.1080/19388079209558078

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Prentice-Hall, Inc.

Reynolds, R. E., & Anderson, R. C. (1982). Influence of questions on the allocation of attention during reading. *Journal of Educational Psychology*, 74(5), 623–632. doi:10.1037/0022-0663.74.5.623

Rickards, J. P., & Denner, P. R. (1979). Depressive effects of underlining and adjunct questions on children's recall of text. *Instructional Science*, *8*(1), 81-90. doi: 10.1007/BF00054983

Ryan, T. E. (2006). Motivating novice students to read their textbooks. *Journal of Instructional psychology*, 33(2).

Salmerón, L., & Llorens, A. (2019). Instruction of digital reading strategies based on eye-movements modeling examples. *Journal of Educational Computing Research*, 57(2), 343-359.

Saumell, L., Hughes, M. T., & Lopate, K. (1999). Underprepared college students' perceptions of reading: Are their perceptions different than other students?. *Journal of College Reading and Learning*, 29(2), 123-135.

Schaffner, E., & Schiefele, U. (2013). The prediction of reading comprehension by cognitive and motivational factors: Does text accessibility during comprehension testing make a difference?. *Learning and Individual Differences*, 26, 42-54.

Taraban, R., Rynearson, K., & Kerr, M. (2000). College students' academic performance and self-reports of comprehension strategy use. *Reading Psychology*, 21(4), 283-308.

van der Schoot, M., Reijntjes, A., & van Lieshout, E. C. (2012). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing*, 25(7), 1665-1690.

Van Der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal of Research in Reading*, 31(2), 203-223.

Vidal-Abarca, E., Mañá, A., & Gil, L. (2010). Individual differences for self-regulating task-oriented reading activities. *Journal of Educational Psychology*, 102(4), 817.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology*, 27(4), 551-572.

Yeari, M., & Lantin, S. (2021). The origin of centrality deficit in text memory and comprehension by poor comprehenders: a think-aloud study. *Reading and Writing*, *34*(3), 595-625.

Yeari, M., Oudega, M., & van den Broek, P. (2017). The effect of highlighting on processing and memory of central and peripheral text information: Evidence from eye movements. *Journal of Research in Reading*, *40*(4), 365-383. doi: 10.1111/1467-9817.12072

Yeari, M., van den Broek, P., & Oudega, M. (2015). Processing and memory of central versus peripheral information as a function of reading goals: Evidence from eye-movements. *Reading and writing*, *28*(8), 1071-1097. doi: 10.1007/s11145-015-9561-4

Yue, C. L., Storm, B. C., Kornell, N., & Bjork, E. L. (2015). Highlighting and its relation to distributed study and students' metacognitive beliefs. *Educational Psychology Review*, *27*(1), 69-78. doi: 10.1007/s10648-014-9277-z

Zargar, E., Adams, A. M., & Connor, C. M. (2020). The relations between children's comprehension monitoring and their reading comprehension and vocabulary knowledge: an eye-movement study. *Reading and writing*, *33*(3), 511-545. doi: 10.1007/s11145-019-09966-3