Analysis of the Readability of Health Education Texts for Elderly Readers: An Eye-Tracking Experiment

Hui-Xin Chen, National Taiwan University of Science and Technology, Taiwan
Hou-Chiang Tseng, National Taiwan University of Science and Technology, Taiwan
Yao-Ting Sung, National Taiwan Normal University, Taiwan

The Asian Conference on Education 2024
Official Conference Proceedings

Abstract

Health literacy refers to the ability to understand, evaluate, and apply health information. Individuals with higher health literacy are able to comprehend the content of health education materials. In contrast, those with lower health literacy may struggle to understand the information, potentially leading to worsening health conditions. As the global population aged 65 and above continues to grow rapidly and cognitive abilities decline, providing more readable texts can enhance reading comprehension. Thus, assessing the readability of health education texts has become an important research topic. To improve reading comprehension, different countries have established their readability guidelines. For example, the United States recommends texts suitable for a 5th to 6th-grade reading level. However, Taiwan currently lacks similar guidelines for health education texts. Given this research gap, the present study employs eye-tracking experiments to gather physiological data from the reader's perspective. This approach helps verify the reading process and comprehension performance, ensuring that the content is effectively understood. In the analysis, the eye-tracking data from 11 participants were used, focusing on five commonly studied eye-tracking indicators. The results showed that the regression in count for diabetes proper nouns reached a significant level. This study recommends that the difficulty of health education texts should not exceed a 6th-grade reading level.

Keywords: Elderly, Education Materials, Reading Ability, Readability, Eye Movements

1. Introduction

With the continuous advancement of medical technology, health literacy has increasingly become an important evaluation indicator. Health literacy refers to an individual's ability to understand, evaluate, and apply health information in daily life (Sørensen et al., 2012). Individuals with high health literacy can quickly comprehend health information, which helps to enhance their personal capabilities and eliminate doubts regarding health-related information (Kickbusch et al., 2013; Rudd, 2013). For example, during the COVID-19 pandemic, healthcare institutions commonly used online methods to disseminate information. However, the content presented in these texts was often too complex for the general public, leading to a lack of full comprehension and missed opportunities for self-protection (Szmuda et al., 2020). It can be concluded that providing easily understandable text content can effectively address these issues and help individuals find more suitable treatment options (Brown et al., 2004; Howell et al., 2017).

In fact, even without the pandemic, the readability of texts is crucial for older adults. With the growing proportion of the global population aged 65 and above, the demand for healthcare services has also increased (McNicoll, 2002). For example, there is greater emphasis on oral health care and hearing care (Lee et al., 2020; Wallhagen & Strawbridge, 2017). Providing written materials or pamphlets can assist older adults in recalling important information and allow them to learn at their own pace (Bernier, 1993). Goodman and Lambert (2023) and Pearman and Storandt (2005) further pointed out that the decline in cognitive function among older adults should be considered when assessing their ability to understand texts. Providing easily understandable medication information for older adults can help reduce side effects caused by improper medication use (Liu et al., 2014).

To provide suitable reading materials, the American Medical Association (AMA) and the U.S. Department of Health and Human Services (USDHHS) recommend that the difficulty of texts should be kept at the level of 5th to 6th grade (Kutner et al., 2006). In the UK, guidelines recommend that health education materials for older adults should not exceed a readability level of 12 (Petterson, 1994), corresponding to the 11th-grade level in the U.S. (Gunning, 1952). In Australia, the SA Health agency recommends that texts should be at an 8th-grade reading level to ensure comprehension (Cheng & Dunn, 2015).

In summary, Western countries have established guidelines for the grade level at which texts should be understandable. Whether on the website or on paper, there are standards that can be referenced. In contrast, there is relatively little research in Taiwan that explores the appropriate grade level for health education texts. Therefore, the purpose of this study is to explore the level of difficulty in health education texts that are most suitable for older adults to read. Eye-tracking experiments will be conducted from the readers' perspective to understand their reading process and reading comprehension levels.

2. Literature Review

Readability is the extent to which readers can understand the text (Dale & Chall, 1949; Klare, 1963, 2000; McLaughlin, 1969). Readable text allows readers to understand and absorb content more effectively (DuBay, 2007). For example, using easily understood vocabulary, simpler sentence structures, and appropriate article length can enhance readability (Klare, 2000; Van Den Broek & Kremer, 2000). Therefore, when evaluating the readability of a text, it is essential to consider the reader’s ability and strive to minimize cognitive load.

In previous research, readability assessments typically employed readability formulas. Especially in English-speaking countries, these formulas provide a standardized quantitative metric, allowing the difficulty of a text to be measured and evaluated directly (DuBay, 2007). Traditional readability formulas are primarily used to assess the complexity of sentence or word structures in a text (Pruthi et al., 2015). By calculating indicators such as word length, sentence length, and the proportion of difficult words, the readability grade level of the text can be determined. This value indicates the appropriate grade level for reading the text. Common readability formulas are shown in Table 1.

In addition to using readability formulas to assess text difficulty, some researchers have employed eye-tracking experiments to confirm participants' reading processes and reading comprehension strategies. During these experiments, eye-tracking data offer insights into what readers focus on, the order in which they read the text, and the time they spend on different sections (Holsanova et al., 2009). In addition, eye-tracking experiments can be used to identify individual differences between readers. For example, students with adequate prior knowledge tend to experience a lower cognitive load when learning specific domain content. Therefore, this results in better learning outcomes (Jarodzka et al., 2010). Thus, eye-tracking experiments can help us understand differences in text comprehension among readers with varying levels of reading ability.

To gain deeper insights into participants' comprehension patterns, eye-tracking indicators can be subjected to analysis. These indicators derived from different types of eye movements, include Total Fixation Duration (TFD), Total Fixation Count (TFC), Percentage of Fixation Duration, First Fixation Duration, Number of regressive saccades, and so forth. Common eye-tracking indicators are detailed in Table 2. By analyzing several common eye-tracking indicators, the aim is to understand the reading comprehension and reading performance of older adults and to provide them with appropriate texts to read.

Table 1: Common Readability Formulas and Indicators

Formula name	Calculation formula	Indicators
Flesch Reading Ease (Flesch, 1948; Flesch, 1979)	Reading Ease = 206.835 – (1.015 × ASL) – (84.6 × ASW)	Average Sentence Length、Average Number of Syllables per Word
Gunning FOG (Gunning, 1952)	Gunning Fog Index = (ASL + PHW) ×0.4	Average Sentence Length、Percentage of Hard Words
SMOG (Mc Laughlin, 1969)	$\text{SMOG Grade} = 1.0430 \times \sqrt{\text{number of polysyllables} \times (\frac{30}{\text{number of sentences}})} + 3.1291$	Number of polysyllables、Number of sentenced

Table 2: Common Eye-Tracking Indicators

Eye-tracking Indicators	Definition/Measurement	References
Total fixation duration	The total time spent on all fixation points, with longer durations indicates a greater cognitive processing load.	Hannus & Hyönä, (1999); Hegarty & Just, (1993)
Total fixation count / the number of fixations	The number of fixations within all AOI. Areas with more fixations require more time for cognitive processing.	Eitel, (2016); Schnotz & Wagner, (2018)
Percentage of fixation duration	The fixation duration within the AOI / the total fixation duration on the screen. It reflects the reader's attention allocation to each text segment.	Alemdag & Cagiltay, (2018)
First fixation duration	The duration of the reader's first fixation on a word reflects the degree of selective attention given to the word.	Alemdag & Cagiltay, (2018); Lai et al., (2013); Scheiter & Eitel, (2015)
Number of regressive saccades	It is the total number of times the eyeball jumps from the position of the rear text to the front, reflecting the reader's late processing of single words. When readers encounter words or sentences they don't understand, they will look back more often.	Jian et al., (2013); Mason et al., (2016)

3. Method

This study conducts eye-tracking experiments focused on older adults. The objective of this study is to examine how text difficulty influences the reading comprehension of older adults with varying reading abilities. Eye-tracking indicators can reflect readers' cognitive processing and learning performance of text. For example: dwell time, total fixation count, dwell time %, first fixation duration, and regression in count. In addition, by asking participants to complete the Diabetes Knowledge Questionnaire (DKQ), it can be further investigated whether eye movement performance is related to the knowledge background of the reader or is caused by differences in individual reading ability.

3.1 Participants

For this study, 15 participants aged 65 and older were recruited. They lived in Taipei City. All participants were native Chinese speakers, and it was confirmed that their vision was either normal or corrected to a level that would be considered normal. Four older adults were excluded from the study due to not meeting the age criteria or failing to record eye-tracking data. Ultimately, 11 participants were included in the analysis.

3.2 Apparatus

The experiment used the SR Research Eyelink 1000, which recorded eye movements at a sampling rate of 1000 Hz. A chin bar was used to stabilize the participants' heads. The reading text was presented on a 24-inch LCD monitor with a resolution of 1920 x 1200 pixels. The text size was set to 28, and the distance between the monitor and the participants was 65 cm.

3.3 Experimental Materials

This experiment focused on the topic of diabetes due to the continuous global increase in the number of diabetes patients, particularly among older adults. A review of the statistical data shows that the worldwide prevalence of diabetes is expected to rise to 642 million by 2040, with the largest increase projected among individuals aged 60 to 79 (Ogurtsova et al., 2017). In Taiwan, the incidence of diabetes and cancer is also on the rise. According to the 2023 statistical data from the Ministry of Health and Welfare on the ten leading causes of death, malignant neoplasms (cancer) ranked first, accounting for 25.8% of total deaths, while diabetes was ranked fifth. The primary causes of death among diabetes patients are closely related to malignant tumors, cardiovascular diseases, and other conditions. Notably, 87% of cancer-related deaths occur in individuals aged 55 and above. Providing health education texts that are suitable for reading can significantly reduce the cognitive load of the elderly, help them understand the disease more deeply, and take appropriate measures according to the symptoms and risks (Conner et al., 2019; Ebaid & Crewther, 2019). In light of this, the study selected diabetes as the theme and categorized the texts read by participants into three levels of difficulty: easy (6th grade), medium (9th grade), and Hard (12th grade). The texts were authored by six professional teachers, verified for readability, and reviewed by professional doctors to ensure content accuracy.

3.4 Reading Ability Test

Given the dearth of knowledge regarding the range of reading abilities among Taiwanese elders, this study employed the Diagnostic Assessment of Chinese Competence (DACC) to gain insight into the comprehension abilities of older adults across varying reading ability levels. Through computerized adaptive tests, participants' reading abilities are assessed in five aspects: word recognition, superficial meaning comprehension, textual meaning integration, inferential comprehension, and analysis and evaluation. The test results showed that the participants’ reading abilities ranged from second to twelfth grade (Lee et al., 2021). Based on the DACC test results, this study categorized participants into two groups for further analysis: low reading comprehension (below 7th grade) and high reading comprehension (above 8th grade).

3.5 Diabetes Knowledge Questionnaire

To determine whether differences in participants' eye movement behaviors stem from their background knowledge of diabetes or their reading abilities, this study first had participants complete the Diabetes Knowledge Questionnaire (DKQ). This questionnaire assesses their understanding of diabetes knowledge and serves as a reference indicator during the reading experiments. The DKQ is a commonly used tool for measuring diabetes knowledge, with good internal consistency, indicated by a Cronbach's α coefficient of .78 (Garcia et al., 2001). A staff member with a bachelor's degree in English translated the DKQ, which was subsequently reviewed by two experts: a psychometrician and a physician certified in diabetes education. The translation process also referenced the simplified Chinese version developed by Hu and colleagues (2013), which reported a Cronbach's α coefficient of .89. The questionnaire consists of 24 brief statements, with each item offering three response options: true, false, or unsure. The total score ranges from 0 to 24, with higher scores indicating a greater understanding of diabetes knowledge among the participants.

According to the results in Table 3, individuals with low reading comprehension abilities had an average score of 15.00 on the questionnaire, while those with high reading comprehension abilities had an average score of 16.67. Individuals with high reading comprehension abilities performed better than those with low reading comprehension abilities. Additionally, the standard deviation for high reading comprehension participants is 2.73, which is higher than the 2.00 for low reading comprehension participants. The t-value is -1.13, with a p-value of .29, indicating that the results are not statistically significant.

Reading comprehension is closely linked to vocabulary mastery and reading fluency (Ehri, 2014). Individuals with better reading comprehension are generally more skilled at understanding the vocabulary within the text, which helps them integrate and infer the content while adapting their reading strategies (Cain et al., 2004; Connor et al., 2015; Zargar et al., 2020). Moreover, readers with relevant background knowledge can better synthesize the meanings of sentences, thereby enhancing their reading comprehension skills (Mayer, 2005; Perfetti & Stafura, 2014; Schnotz & Bannert, 2003; Schnotz et al., 2014). Therefore, to investigate whether eye movement behavior is influenced by the reader's background knowledge, this study employs independent samples t-tests to examine the relationship between participants' performance on the knowledge questionnaire and their reading comprehension across different reading abilities. In statistics, independent samples t-tests are commonly used to explore differences between two distinct groups, such as comparing the performance of experts and novices in solving math problems (Chen & Wen, 2023). However, according to the results, both groups have a similar level of diabetes knowledge, which does not significantly affect their comprehension of diabetes-related texts.

Since the scores on the diabetes knowledge questionnaire did not show statistically significant differences, the subsequent eye-tracking analysis excluded the influence of diabetes background knowledge. This allows for an exploration of whether there are differences in eye movement comprehension performance among participants with varying reading abilities. The analysis uses eye-tracking indicators to examine the cognitive processes of participants with different reading abilities as they read texts of varying difficulty, thereby facilitating further inferences.

Table 3: t-Test Analysis of Correct Responses on the Questionnaire by Reading Ability

	n	M	SD	t
Low reading ability	5	15.00	2.00	
High reading ability	6	16.67	2.73	-1.13

$*p < .05$; $**p < .01$

3.6 Procedure

The experimental flow chart of this study is shown in Figure 1. Before the experiment officially begins, the contents of the informed consent form are first explained to the participants. After the participants understood the entire experimental process, they signed a consent form to express their agreement to participate in the experiment. Subsequently, they completed the DKQ to assess their prior knowledge of diabetes. Following this, the DACC was conducted to evaluate the participants' reading abilities. The duration of the testing phase was approximately 40 minutes to 1.5 hours, after which participants were permitted a 10-minute break before commencing the eye-tracking experiment.

Before officially starting the eye-tracking experiment, each participant underwent a nine-point calibration and verification process to adjust the camera's position and focus, ensuring that their pupils were aligned with the central cross on the screen. The eye-tracking experiment was then initiated and divided into three stages. After reading each page of text, participants pressed the space bar to proceed to the next page. When the screen displays "Congratulations, you have completed this section", participants are asked to answer questions about the text's topic. After responding, participants were given a 5 to 10-minute break. The entire eye-tracking experiment lasted approximately 20 minutes.

For subsequent analyses, this study utilized heatmaps to evaluate participants' comprehension of the text and their reading behaviors. This involved calculating the fixation duration and frequency in areas related to diabetes proper nouns and non-diabetes proper nouns. The rationale for focusing on these two main categories is that vocabulary significantly impacts readers' understanding of the text. According to Chaffin et al. (2001), readers usually spend more time on cognitive processing when they encounter novel or unfamiliar words. Jian et al. (2013) also explored how readers understand academic vocabulary in physics, finding that readers can infer the meanings of these terms through contextual cues and other familiar vocabulary. Following this, a quantitative analysis of Area of Interest (AOI) was conducted to extract five common eye-tracking indicators, aiming to gain deeper insights into the participants' reading comprehension processes.

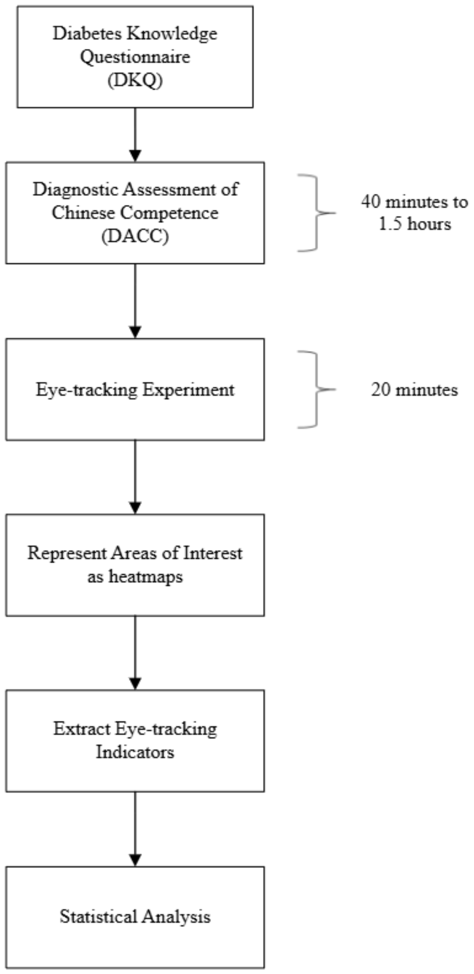

Figure 1: The Experimental Flow Chart

4. Results

This study used heatmaps to visually analyze eye-tracking data, giving an initial look at how participants with different reading skills performed when reading texts of different difficulty levels. Figures 2, 3, and 4 illustrate the findings. The heatmaps show text paragraphs marked with varying color intensities, with darker colors indicating longer reading times and more frequent fixations. Following this initial analysis, diabetes proper nouns and non-diabetes proper nouns were identified as Area of Interest (AOI). The diabetes proper nouns were selected based on definitions provided by the International Diabetes Federation (2024), the American Diabetes Association (2024), and SA Health (2022). Non-diabetes proper nouns were excluded from the specialized terms vocabulary. Subsequent eye-tracking analysis was then conducted on these AOI, as depicted in Figures 5 and 6.

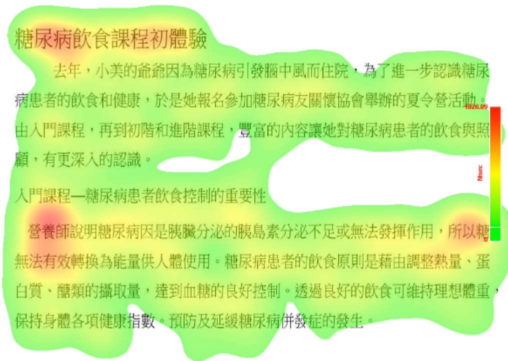

Figure 2: Heatmap of Low Reading Ability Participants Reading a Text of Easy Difficulty

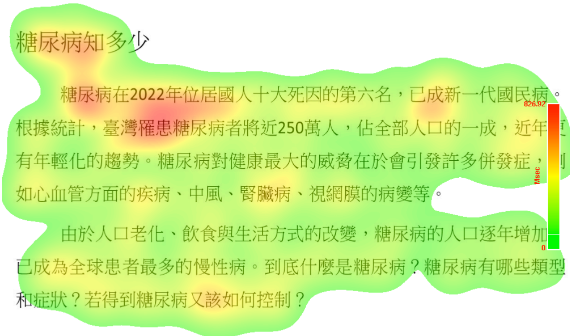

Figure 3: Heatmap of Low Reading Ability Participants Reading a Text of Medium Difficulty

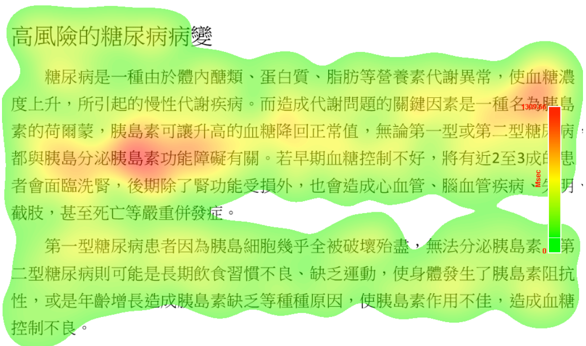

Figure 4: Heatmap of Low Reading Ability Participants Reading a Text of Hard Difficulty

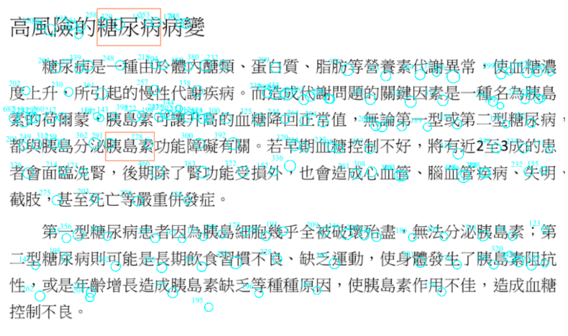

Figure 5: Heatmap With Diabetes Proper Nouns As AOI

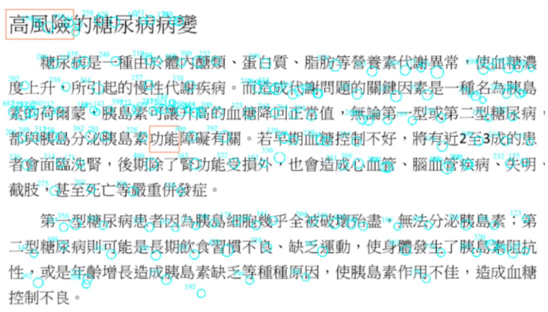

Figure 6: Heatmap With Non-diabetes Proper Nouns As AOI

4.1 Data Selection and Analysis

According to Rayner (2009), participants' gaze duration typically ranges between 100 and 500 milliseconds. If the gaze duration exceeds 1,000 milliseconds, it may be attributed to instrumental error. Therefore, during the eye-tracking data analysis, data with gaze durations shorter than 100 milliseconds or longer than 1,000 milliseconds were excluded (Morrison, 1984; Rayner & Pollatsek, 2016). Ultimately, data from 11 participants were collected, including 5 participants with low reading ability and 6 participants with high reading ability. For the data analysis, diabetes proper nouns and non-diabetes proper nouns were designated as AOI. The eye-tracking indicators employed were based on commonly used eye-tracking indicators from past studies (Jian & Ko, 2017; Jian et al., 2013; Rayner, 1998; Rayner et al., 2006; Schad et al., 2014), such as dwell time, total fixation count, dwell time %, first fixation duration and regression in count. Commonly used eye-tracking indicators are detailed in Table 4.

Table 4: Common Eye-Tracking Indicators

Eye-tracking Indicators	Definition/Measurement	References
Dwell time	The total time spent on all fixation points within the AOI is summed, with longer durations indicating a greater cognitive processing load.	Jian & Ko, (2017); Lai et al., (2013)
Fixation count	The total number of fixations across all AOIs is summed, with a higher number of fixations indicating a higher degree of cognitive processing or that the text information is more engaging.	Eitel, (2016); Schnotz & Wagner, (2018)
Dwell time %	The proportion of dwell time on the AOI relative to the total dwell time on the entire text reflects the reader's selective attention allocation or the time spent on processing the information within the AOI.	Alemdag & Cagiltay, (2018); Mason et al., (2013)
First fixation duration	The first fixation duration reflects the reader's initial semantic processing or the degree of selective attention given to a word; the more attractive or familiar the word is to the reader, the shorter the fixation duration will be.	Alemdag & Cagiltay, (2018); Jian et al., (2013); Scheiter & Eitel, (2015)

Regression in count	The total number of regressions, where the eyes jump backward to earlier positions in the text, reflects the reader's late-stage processing of words. When encountering unfamiliar words or sentences, the number of regressions increases.	Jian et al., (2013); Mason et al., (2016)

4.2 Statistical Analysis Methods

In the statistical analysis, this study followed the experimental design of Jian and Ko (2017), using reading ability (low and high) and text difficulty (medium and hard) as independent variables, with eye-tracking measures as the dependent variables. During the experiment, participants read two articles, and data analysis was based on eye-tracking indicators selected from the AOI. Additionally, following the study by Jou and Mariñas (2023), which assessed the reading behaviors of individuals with dyslexia, the independent variables were set as people without dyslexia and those with dyslexia, as well as three different text designs, with eye-tracking measures again as the dependent variables. Participants read three different text designs in sequence, and subsequent analysis and inference were conducted based on the eye-tracking indicators.

Upon reviewing the two studies mentioned above, a common approach can be observed: the use of a two-factor mixed design, where reading ability is treated as a between-subjects factor and text difficulty as a within-subjects factor, with various eye-tracking indicators as dependent variables. This aligns with the design approach of this study. Therefore, in this study, the independent variables were set as low reading ability (below 7th grade), high reading ability (above 8th grade), and text difficulty (6th, 9th, and 12th-grade levels). The dependent variables were five eye-tracking indicators, aiming to explore the performance differences of participants with different reading abilities when reading texts of varying difficulty, and to infer their eye-tracking behavior further.

4.2.1 Analysis Results Using Diabetes Proper Nouns and Non-diabetes Proper Nouns As Units

In selecting the AOI, the analysis primarily focused on diabetes proper nouns and non-diabetes proper nouns. The descriptive statistics summary is presented in Table 5, and the ANOVA summary table is shown in Table 6. Among the five eye-tracking indicators, only the regression in count for diabetes proper nouns reached significant interaction between reading ability and text difficulty, $F(2, 18) = 3.69$, $p < .05$, $\eta^2 = .29$. This eye-tracking indicator may reflect eye movement behavior when readers have difficulty understanding the text or vocabulary, leading them to integrate or clarify their perspective (Rayner et al., 2003). Additionally, text difficulty also exhibited a significant main effect, $F(2, 18) = 5.75$, $p < .05$, $\eta^2 = .39$.

From the results in Table 7, it can be seen that the simple main effect of reading ability significantly impacted regression in count for easy texts, $F(1, 27) = 5.09$, $p < .01$, $\eta^2 = .16$. Post-hoc comparisons indicated that participants with low reading comprehension ($M = 0.87$, $SD = 0.51$) exhibited significantly higher regression counts when reading easier texts than those with high reading comprehension ($M = 0.42$, $SD = 0.25$). Research suggests that participants with low reading comprehension tend to learn from simpler text and images primarily because easier texts are more comprehensible for them. Consequently, they concentrate more attention on these simpler texts, leading to a significant increase in the number of fixations and total reading time.

The analysis of regression behaviors also indicated that participants with low reading comprehension had significantly more regressions in easier texts compared to more difficult ones. In other words, individuals with low reading comprehension tend to spend more time and focus on easily understandable texts, aiding their learning and information integration within these materials (Jian & Ko, 2017). However, this study categorized text difficulty into three levels (6th, 9th, and 12th grade) and classified participants' reading comprehension into low (below 7th grade) and high (above 8th grade). As the reading ability of the participants in this study ranged from 5th to 7th grade, providing 6th-grade level texts does not create a significant cognitive gap. In summary, the complexity of the health education texts should be kept at a level that is easily understandable for 6th graders.

On the other hand, the results in Table 7 indicate that the simple main effect of text difficulty significantly influenced participants with high reading comprehension, $F(2, 18) = 10.16$, $p < .01$, $\eta^2 = .06$. Post-hoc comparisons revealed that participants had higher regression in counts when reading texts of medium difficulty ($M = 1.25$, $SD = 0.29$) and high difficulty ($M = 1.14$, $SD = 0.34$) compared to easier texts ($M = 0.42$, $SD = 0.25$). This suggests that even readers with good comprehension skills may need to put in extra effort to understand content when reading texts that match their reading level.

This phenomenon is similar to what is seen in individuals with low reading comprehension, who often take more time to understand texts. As text difficulty increases, readers' fixation durations and regression counts also rise (Jacobson & Dodwell, 1979; Rayner et al., 1989; Rayner et al., 2006). An increase in fixation count reflects the reader's need for higher-level cognitive processing when comprehending the text (Eitel, 2016; Schnotz & Wagner, 2018). Therefore, participants with a high level of reading comprehension tend to have significantly more regressions when they read texts of medium to high difficulty than when they read texts of lower difficulty.

Table 5: Descriptive Statistics Summary of Regression in Count by Reading Ability and Text Difficulty (diabetes proper nouns)

Text Easy		Text Medium		Text Hard	
Low reading ability	High reading ability	Low reading ability	High reading ability	Low reading ability	High reading ability
5	6	5	6	5	6
0.87	0.42	1.00	1.25	0.89	1.14
0.51	0.25	0.30	0.29	0.30	0.34

Table 6: ANOVA Summary Table for Regression in Count by Reading Ability and Text Difficulty (diabetes proper nouns)

SV	SS	df	MS	F	η2
Reading Ability (A)	0.00	1	0.00	0.02	.00
Error(A)	0.89	9	0.10		
Text difficulty (B)	1.40	2	0.70	5.75*	.39
Reading Ability (A) x Text difficulty (B)	0.90	2	0.45	3.69*	.29
Error	2.19	18	0.12		

*$p < .05$

Table 7: Summary Table of Simple Main Effects for Regression in Count by Reading Ability and Text Difficulty (diabetes proper nouns)

SV	*SS*	*df*	*MS*	*F*	η2	Post hoc Comparison
Reading Ability (A)						
Text Easy (b1)	0.56	1	0.56	5.09*	.16	Low>High
Text Medium (b2)	0.18	1	0.18	1.64	.06	
Text Hard (b3)	0.16	1	0.16	1.45	.05	
Error	3.08	27	0.11			
Text difficulty (B)						
Low reading ability (a1)	0.05	2	0.02	0.19	.02	
High reading ability (a2)	2.47	2	1.23	10.16*	.53	Medium, Hard>Easy
Error	2.19	18	0.12			

*p<.01

5. Conclusion

Previous studies primarily focused on elementary and university students to provide appropriate reading materials for them. In contrast, this research focuses on older adults as the participants, whose cognitive comprehension differs from students. To understand the reading abilities of older adults and provide suitable health education texts, this study employed testing and eye movement experiments to observe the reading behaviors and performances of the participants.

In selecting the AOI, the analysis concentrated on diabetes proper nouns and non-diabetes proper nouns, utilizing five eye-tracking indicators. The results revealed a significant effect solely in the regression in count for diabetes proper nouns. Specifically, participants with low reading comprehension showed more regressions when reading simpler texts than those with high reading comprehension; this indicates that individuals with lower reading comprehension may be more willing to spend additional time understanding easier texts. Moreover, participants with high reading comprehension exhibited significantly higher regression in counts when reading medium and hard texts than when reading easy texts. This suggests that even highly proficient readers put in significant effort to understand texts that match their reading level, resulting in longer fixation durations and more regressions as the text difficulty increases. This shows that although cognitive decline does occur in older adults, the results are consistent with previous research and help to fill in the gap. In summary, the 11 participants in this study were from the Taipei metropolitan area. Five of them were identified as having low reading comprehension skills. This group included one fifth grader, one sixth grader, and three seventh graders. In theory, all of these participants have relatively good reading comprehension skills. However, in order to take into account the reading comprehension levels of older adults in non-metropolitan areas, we recommend that the level of difficulty of the text provided should be no higher than the sixth grade level.

6. Limitations of the Study

This study focused on five common eye-tracking indicators for analysis. However, the range of available eye-tracking indicators is extensive. Future research could consider incorporating additional indicators from different dimensions, such as average saccade length and fixation position as spatial metrics. This would provide a more comprehensive observation of participants' eye movement behaviors and enable deeper exploration. In addition, this study had limitations related to sample size, which may affect the results of statistical analyses. To enhance the inferential strength of future research, it is advisable to increase the sample size and collect more data for a more thorough statistical evaluation.

7. Acknowledgments

This project was financially supported by the Empower Vocational Education Research Center of National Taiwan University of Science and Technology (NTUST) from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan. This project was also supported by funding from the National Science and Technology Council, Taiwan (111-2410-H-011 -039 -MY2).

References

Ab Hamid, A. R., Hashim, M. F., Hasan, N. A., & Azhan, N. M. (2020). Readability of COVID-19 information by the Malaysian Ministry of Health. *Journal of Nusantara Studies (JONUS)*, *5*(2), 170-191.

Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, *125*, 413-428.

American Diabetes Association. (n.d.). Common Terms. https://diabetes.org/about-diabetes/common-terms (Accessed September 2024)

Bernier, M. J. (1993). Developing and evaluating printed education materials: a prescriptive model for quality. *Orthopaedic Nursing*, *12*(6), 39-46.

Brown, R. F., Butow, P. N., Butt, D. G., Moore, A. R., & Tattersall, M. H. (2004). Developing ethical strategies to assist oncologists in seeking informed consent to cancer clinical trials. *Social science & medicine*, *58*(2), 379-390.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31.

Castro-Sánchez, E., Spanoudakis, E., & Holmes, A. H. (2015). Readability of Ebola information on websites of public health agencies, United States, United Kingdom, Canada, Australia, and Europe. *Emerging infectious diseases*, *21*(7), 1217.

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: a study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 225.

Chen, W.-Y., & Wen, M. L. (2023). Exploring Differences Between Expert and Novice in Mathematical Proof Validation Using Eye-Tracking Technology. International Journal on Digital Learning Technology, 15(2), 33-58.

Cheng, C., & Dunn, M. (2015). Health literacy and the Internet: a study on the readability of Australian online health information. *Australian and New Zealand journal of public health*, *39*(4), 309-314.

Conner, L. R., Fernández, Y., Junious, E., Piper, C., & Rowan, D. (2019). Evaluating HIV educational materials for older people. *Journal of the International Association of Providers of AIDS Care (JIAPAC)*, *18*, 2325958219849054.

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.

Dale, E., & Chall, J. S. (1949). The concept of readability. *Elementary English*, *26*(1), 19-26.

DuBay, W. H. (2007). *Smart Language: Readers, Readability, and the Grading of Text*. ERIC.

Ebaid, D., & Crewther, S. G. (2019). Visual information processing in young and older adults. *Frontiers in aging neuroscience*, *11*, 449620.

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, *18*(1), 5-21.

Eitel, A. (2016). How repeated studying and testing affects multimedia learning: Evidence for adaptation to task demands. *Learning and Instruction*, *41*, 70-84.

Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, *32*(3), 221.

Flesch, R. (1979). How to write plain English: Let's start with the formula. *University of Canterbury*.

Garcia, A. A., Villagomez, E. T., Brown, S. A., Kouzekanani, K., & Hanis, C. L. (2001). The Starr County Diabetes Education Study: development of the Spanish-language diabetes knowledge questionnaire. *Diabetes care*, *24*(1), 16-21.

Goodman, C., & Lambert, K. (2023). Scoping review of the preferences of older adults for patient education materials. *Patient Education and Counseling*, *108*, 107591.

Gunning, R. (1952). The technique of clear writing. *(No Title)*.

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low-and high-ability children. *Contemporary educational psychology*, *24*(2), 95-123.

Hazawawi, N. A. M., Zakaria, M. H., & Hisham, S. (2016). SPIKE: Online reading competencies measure for Malay language. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *8*(2), 123-127.

Hegarty, M., & Just, M.-A. (1993). Constructing mental models of machines from text and diagrams. *Journal of memory and language*, *32*(6), 717-742.

Holsanova, J., Holmberg, N., & Holmqvist, K. (2009). Reading information graphics: The role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(9), 1215-1226.

Howell, D., Harth, T., Brown, J., Bennett, C., & Boyko, S. (2017). Self-management education interventions for patients with cancer: a systematic review. *Supportive Care in Cancer*, *25*, 1323-1355.

Hu, J., Gruber, K. J., Liu, H., Zhao, H., & Garcia, A. A. (2013). Diabetes knowledge among older adults with diabetes in Beijing, China. *Journal of clinical nursing*, *22*(1-2), 51-60.

International Diabetes Federation. (n.d.). Diabetes Glossary. https://idf.org/diabetes-glossary/(Accessed September 2024)

Jacobson, J. Z., & Dodwell, P. C. (1979). Saccadic eye movements during reading. *Brain and Language*, *8*(3), 303-314.

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction*, *20*(2), 146-154.

Jian, Y.-C., & Ko, H.-W. (2017). Influences of text difficulty and reading ability on learning illustrated science texts for children: An eye movement study. *Computers & Education*, *113*, 263-279.

Jian, Y. C., Chen, M. L., & Ko, H. w. (2013). Context Effects in Processing of Chinese Academic Words: An Eye-Tracking Investigation. *Reading Research Quarterly*, *48*(4), 403-413.

Jou, Y.-T., & Mariñas, K. A. A. (2023). Developing inclusive lateral layouts for students with dyslexia–Chinese reading materials as an example. *Research in Developmental Disabilities*, *132*, 104389.

Kickbusch, I., Pelikan, J. M., Apfel, F., & Tsouros, A. D. (2013). Health literacy: The solid facts World Health Organization. Regional Office for Europe.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel.

Klare, G. (1963). The measurement of readability. In: Iowa State University Press.

Klare, G. R. (2000). The measurement of readability: useful information for communicators. *ACM Journal of Computer Documentation (JCD)*, *24*(3), 107-121.

Kutner, M., Greenburg, E., Jin, Y., & Paulsen, C. (2006). The Health Literacy of America's Adults: Results from the 2003 National Assessment of Adult Literacy. NCES 2006-483. *National Center for education statistics*.

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., & Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational research review*, *10*, 90-115.

Lee, K. H., Choi, Y. Y., & Jung, E. S. (2020). Effectiveness of an oral health education programme for older adults using a workbook. *Gerodontology*, *37*(4), 374-382.

Lee, Y. H., Chou, Y. T., & Sung, Y. T. (2021). Development and Validation of a Diagnostic Assessment of Chinese Competence System. *Bulletin of Educational Psychology*, *53*(2), 285-306.

Liu, F., Abdul-Hussain, S., Mahboob, S., Rai, V., & Kostrzewski, A. (2014). How useful are medication patient information leaflets to older adults? A content, readability and layout analysis. *International journal of clinical pharmacy*, *36*, 827-834.

Mason, L., Pluchino, P., & Tornatora, M. C. (2013). Effects of picture labeling on science text processing and learning: Evidence from eye movements. *Reading Research Quarterly*, *48*(2), 199-214.

Mason, L., Pluchino, P., & Tornatora, M. C. (2016). Using eye□tracking technology as an indirect instruction tool to improve text and picture processing and learning. *British Journal of Educational Technology*, *47*(6), 1083-1095.

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge university press.

McLaughlin, G. H. (1969). SMOG grading-a new readability formula. *Journal of reading*, *12*(8), 639-646.

McNicoll, G. (2002). World Population Ageing 1950-2050. *Population and development Review*, *28*(4), 814-816.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: evidence for parallel programming of saccades. *Journal of Experimental psychology: Human Perception and performance*, *10*(5), 667.

Ogurtsova, K., da Rocha Fernandes, J., Huang, Y., Linnenkamp, U., Guariguata, L., Cho, N. H., Cavan, D., Shaw, J., & Makaroff, L. (2017). IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes research and clinical practice*, *128*, 40-50.

Parker, R., & Ratzan, S. C. (2010). Health literacy: a second decade of distinction for Americans. *Journal of health communication*, *15*(S2), 20-33.

Pearman, A., & Storandt, M. (2005). Self-discipline and self-consciousness predict subjective memory in older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(3), P153-P157.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22-37.

Petterson, T. (1994). How readable are the hospital information leaflets available to elderly patients? *Age and ageing*, *23*(1), 14-16.

Pruthi, A., Nielsen, M. E., Raynor, M. C., Woods, M. E., Wallen, E. M., & Smith, A. B. (2015). Readability of American online patient education materials in urologic oncology: a need for simple communication. *Urology*, *85*(2), 351-356.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457-1506.

Rayner, K., Juhasz, B., Ashby, J., & Clifton Jr, C. (2003). Inhibition of saccade return in reading. *Vision Research*, *43*(9), 1027-1034.

Rayner, K., Murphy, L. A., Henderson, J. M., & Pollatsek, A. (1989). Selective attentional dyslexia. *Cognitive neuropsychology*, *6*(4), 357-378.

Rayner, K., & Pollatsek, A. (2016). Eye movements in reading a tutorial review. *Attention and performance XII*, 327-362.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and aging*, *21*(3), 448.

Rudd, R. E. (2013). Needed action in health literacy. *Journal of Health Psychology*, *18*(8), 1004-1010.

SA Health. (2022). Common Diabetes Terms. https://www.sahealth.sa.gov.au/wps/wcm/connect/public+content/sa+health+internet/conditions/diabetes/common+diabetes+terms

Schad, D. J., Risse, S., Slattery, T., & Rayner, K. (2014). Word frequency in fast priming: Evidence for immediate cognitive control of eye movements during reading. *Visual cognition*, *22*(3-4), 390-414.

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, *36*, 11-26.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*(2), 141-156.

Schnotz, W., Ludewig, U., Ullrich, M., Horz, H., McElvany, N., & Baumert, J. (2014). Strategy shifts during learning from texts and pictures. *Journal of Educational Psychology*, *106*(4), 974.

Schnotz, W., & Wagner, I. (2018). Construction and elaboration of mental models through strategic conjoint processing of text and pictures. *Journal of Educational Psychology*, *110*(6), 850.

Sørensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., Brand, H., & European, C. H. L. P. (2012). Health literacy and public health: a systematic review and integration of definitions and models. *BMC public health*, *12*, 1-13.

Szmuda, T., Özdemir, C., Ali, S., Singh, A., Syed, M. T., & Słoniewski, P. (2020). Readability of online patient education material for the novel coronavirus disease (COVID-19): a cross-sectional health literacy study. *Public health*, *185*, 21-25.

Van Den Broek, P., & Kremer, K. E. (2000). The mind in action: What it means to comprehend during reading. *Reading for meaning: Fostering comprehension in the middle grades*, 1-31.

Wallhagen, M. I., & Strawbridge, W. J. (2017). Hearing loss education for older adults in primary care clinics: Benefits of a concise educational brochure. *Geriatric Nursing*, *38*(6), 527-530.

Yunus, K. R. M. (1982). *An assessment of structural variables in Malay: A readability formula*. University of Miami.

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.

Contact email: M11211013@mail.ntust.edu.tw