

An Eye-tracking Based Study: The Role of Images and Explanatory Texts in Reading Comprehension

Wang Yuanzhe¹, Zheng Zewen²

¹Huaer Zizhu Academy, Shanghai, China, 201109

²Shanghai Qibao Dwight High School, Shanghai, China, 201104

Abstract: This study aims to explore the influence of the illustrative function of images and explanatory texts on people's reading behaviors. Images are commonly used in books, and their function goes beyond mere decoration. Images can significantly improve people's reading comprehension by serving as illustrations. The study utilizes eye-tracking technology to understand how images influence people's reading processes. Simultaneously, the function of explanatory texts for images is also investigated. An eye tracker records people's gazing process, and eye-tracking parameters are used to quantify the gazing behavior. During the experiment, the stimuli are divided into three groups. Group one consists of an article with only texts. Group two includes an article with texts and a relevant image. Group three contains an article with texts and an image accompanied by an explanatory section. Each group has three articles on different topics. The subjects read different types of stimuli and answered corresponding questions. All the reading processes were recorded by the eye-tracker. Total fixation duration, fixation counts, and average fixation duration are used to evaluate the influence. The results indicate that when images are present, people allocate more attention from texts to images and tend to seek help from pictures. Images enhance people's reading capability and reduce cognitive load. However, there is no significant difference between images alone and images accompanied by explanatory sections. The potential reasons for this are discussed. This finding could improve the understanding of images and explanatory texts as illustrations and provide insights into book or article format design.

Keywords: eye-tracking, illustration, explanatory text, gaze, image, reading, attention, article

1. Introduction

Reading is a fundamental skill essential for individuals to succeed in various aspects of life ^[1]. Over the years, numerous researchers have dedicated their efforts to understanding how to enhance people's reading capabilities. The integration of text and images has emerged as a prominent approach, believed by some to be the most effective method for aiding comprehension ^{[2][3]}. Combining textual information with visual representations not only facilitates readers in grasping the content but also offers valuable cues to enhance understanding. Additionally, in the context of complex technical terms, images prove to be valuable tools for explaining intricate concepts and fostering comprehension ^[4].

Nevertheless, the effectiveness of image usage in reading materials is not without debate. Some researchers posit that if images are too intricate or overwhelming, they may actually impede effective reading ^{[5][6]}. Striking a balance between simplicity and complexity in visual aids is, therefore, crucial.

While several studies have explored the impact of various types of images on reading, few have investigated the effect of explanatory sections in images. In real-world scenarios, many images are accompanied by interpretation sections that utilize brief words, phrases, or sentences to clarify essential information. To shed light on whether images with explanatory sections positively impact reading comprehension, this study undertakes a comparative analysis of pure texts, texts with images, and texts with images containing explanatory sections. Furthermore, the study leverages eye-tracking technology to provide objective and quantitative data on reading behavior. By capturing people's gaze and recording corresponding gazing behaviors, eye trackers offer valuable insights into readers' engagement. Analysis of eye-movement parameters will contribute to a deeper understanding of how article design can be optimized to facilitate enhanced reading comprehension.

The investigation of how text and images influence reading behaviors is an essential avenue for enhancing literacy and understanding complex concepts. This research aims to contribute valuable insights to educational practices and content design, ultimately benefiting learners and readers across

various domains. As technology continues to advance, the incorporation of eye-tracking technology into research methodologies opens up new possibilities for refining instructional materials and promoting effective reading strategies. The quest to unravel the dynamics between images, explanatory sections, and reading comprehension is ongoing, as educators and researchers strive to unlock the full potential of visual aids in the learning process.

2. Materials and Methods

2.1. Participants

The study recruited volunteers, comprising a total sample of 24 participants ($M=26.1$, $SD=11.1$). Among the subjects, there were 10 males and 14 females, all randomly selected from Shanghai, China. Prior to the experiment, participants were informed that it involved reading several articles and answering corresponding questions afterward. It is important to note that all participants voluntarily took part in the study. As a gesture of appreciation, desserts were provided as gifts to the participants upon completion of the experiment. Moreover, before commencing the study, each participant signed a consent form to ensure their willingness to participate.

2.2. Stimuli

The articles, as the independent variables, are divided into three types. The first type consists of only texts. The second type includes texts along with an image related to the article. The third type combines texts with an image featuring explanatory sections. These explanatory sections are labeled close to the image and present key information related to the article, such as defining terms or providing additional context. The contents of the articles contain comparatively uncommon terminologies and knowledge, including topics such as turbulence, the El Niño effect, and Baroque art style. By including uncommon terms, the study aims to exclude any influence brought about by participants' familiarity with the subject matter. It is important to note that all the answers to the questions asked in the experiment were present in the articles.

For the analysis of participants' visual attention when viewing the stimuli, Areas of Interest (AOIs) were classified into five categories: A1 - Text, Answer (the answer found in the texts), Question (the question itself), Whole (Text + Question); A2 - Text, Answer (the answer found in the texts), Question (the question itself), Image (the image), Whole (Text + Question + Image); A3 - Text, Answer (the answer found in the texts), Question (the question itself), Image (the image), and Explanation (the explanatory section), Whole (Text + Question + Image + Explanation). To study participants' visual attention, eye tracking parameters such as total fixation duration (TFD), fixation count (FC), and average fixation duration (AFD) were utilized.

2.3. Design and Procedure

The study utilized a between-subjects design, comprising three experimental groups. Each group of participants read texts with different types of visual aids: Group A1 had no pictures, Group A2 had only pictures, and Group A3 had pictures with explanatory texts.

In Group A1, participants were presented with an explanatory picture of El Niño, a pure picture of turbulence, and no picture of Baroque. In Group A2, participants received the textual content of El Niño along with pure pictures, turbulence without pictures, and Baroque with explanatory pictures. Finally, Group A3 participants were given the textual content of El Niño without pictures, turbulence with explanatory pictures, and Baroque with a pure picture.

All participants willingly agreed to participate in the experiment and provided their informed consent by signing a consent form. Before starting the experiment, the researchers collected personal information, such as age and gender, from each participant.

During the experiment, participants were seated in front of a screen that featured an eye-tracker (Tobii 4C) at the bottom. The eye-tracker was connected to a laptop where the experimenters monitored the participants' gaze. Prior to starting the actual experiment, participants completed a calibration process to ensure accurate gaze recording.

Throughout the experiment, participants were instructed to focus on the reading tasks without moving their heads or bodies. Three different topics of articles were automatically displayed on the screen in

sequence. Each stimulus remained on the screen for 30 seconds, and after that, a question related to the article appeared for 20 seconds. Participants were required to answer the questions within the given time frame. Importantly, the article and the picture remained visible during the 20 seconds as the experiment did not aim to examine people's memory. The display sequence of stimuli was kept consistent across all three groups.

By employing this experimental setup, the study aimed to understand how different types of visual aids, i.e., pictures and explanatory texts, impact participants' reading behaviors and attention during the reading tasks.

2.4. Data analysis

To compare the subjects' gazing behaviors in each group and understand the differences among them, the study conducted the following ANOVA analyses on TFD, FC, and AFD: TFD between A1 Texts vs. A2 Texts vs. A3 Texts; FC between A1 Texts vs. A2 Texts vs. A3 Texts; AFD between A1 Whole vs. A2 Whole vs. A3 Whole.

3. Result

3.1. TFD analysis among A1 and A2 and A3

As Table 1, to investigate the influence of images and explanatory sections on people's reading behaviors, ANOVA analyses were conducted on Text among A1, A2, and A3. The results of the single-factor ANOVA indicated a significant difference in the total fixation duration of participants ($F=5.29$, $p<0.01$) when they viewed the articles with no image ($M=18.23$, $SD=6.41$), the articles with an image ($M=13.79$, $SD=6.49$), and the articles with an image plus an explanatory section ($M=12.70$, $SD=6.52$). The order of durations was $A1 > A2 > A3$.

3.2. FC analysis among A1, A2 and A3

As Table 2, to identify how images and explanatory sections influence people's reading behaviors, ANOVA analyses of Text among A1, A2, and A3 were conducted. A single-factor ANOVA test revealed a significant difference ($F=6.08$, $p<0.01$) in participants' fixation counts when gazing at A1 ($M=85.04$, $SD=20.99$), A2 ($M=68.39$, $SD=27.38$), and A3 ($M=60.54$, $SD=25.77$). Specifically, A1 had a significantly larger fixation count than A2, and A2 had a significantly larger fixation count than A3.

3.3. AFD analysis among A1, A2 and A3

As Table 3, to compare participants' cognitive load during reading comprehension among A1, A2, and A3, a single-factor ANOVA was conducted on A1 Whole, A2 Whole, and A3 Whole. The test revealed a significant difference ($F=3.98$, $p<0.05$) in participants' average fixation duration when exposed to no illustration ($M=0.51$, $SD=0.17$), unexplanatory illustration ($M=0.62$, $SD=0.20$), and explanatory illustration ($M=0.67$, $SD=0.23$). Further post-hoc multiple comparisons (LSD) revealed that participants in A1 had a significantly greater AFD ($p<0.01$) than those in the A2 and A3 groups. However, there was no significant difference ($p>0.05$) between the A2 and A3 groups in terms of AFD.

Table 1: The ANOVA analyses results of TFD among A1, A2 and A3

	Baroque art style	El Niño effect	Turbulence
A1	18.86	19.94	16.48
A2	13.60	12.97	15.10
A3	11.50	14.65	11.96

Table 2: The ANOVA analyses results of FC among A1, A2 and A3

	Baroque art style	El Niño effect	Turbulence
A1	88.50	87.75	78.88
A2	68.63	70.00	66.50
A3	56.13	70.25	55.25

Table 3: The ANOVA analyses results of AFD among A1, A2 and A3

	Baroque art style	El Niño effect	Turbulence
A1	0.52	0.44	0.56
A2	0.64	0.50	0.71
A3	0.61	0.62	0.76

4. Discussions

This study aims to investigate the influence and function of images and explanatory sections during reading comprehension. To achieve this, twenty-four participants were recruited to read three groups of articles. Group A1 contained only texts, Group A2 had both texts and images, and Group A3 contained texts and images with explanatory sections. An eye-tracker was used to record participants' gazing data during the experiment.

The results of ANOVA analyses revealed that participants in Group A1 had higher fixation duration and count compared to those in Group A2 and A3, with Group A2 also showing higher fixation duration and count than Group A3. These findings suggest that the presence of images and explanatory sections influenced people's reading patterns, leading them to allocate more attention to images and explanatory texts. When seeking answers, participants tended to rely on images and explanatory sections rather than solely reading texts.

Furthermore, the average fixation duration (AFD) analysis indicated that participants in Groups A2 and A3 had significantly smaller AFDs than those in Group A1 when reading articles with images. This suggests that images effectively improved participants' reading cognition and reduced their cognitive load during reading, which is consistent with previous research results [7][8][9][10]. Previous studies have shown that images in textbooks serve functions such as summarizing, guiding, awakening aesthetic pleasure, and facilitating cultural inheritance, which helps students better understand the content.

However, in Groups A2 and A3, the AFD results did not show a significant difference, indicating that explanatory sections did not effectively reduce participants' cognitive load. Previous research has demonstrated that the close combination of pictures and text can better help students master knowledge and improve learning efficiency [11]. In this study, the results did not reach the same conclusion, possibly due to the placement of the explanatory sections not being close enough to the images. In previous research, explanatory information was directly marked in the corresponding position on the image, while in this study, the information was placed on the left or right sides, potentially reducing its effectiveness. Future studies should consider directly labeling or marking explanatory information on corresponding positions on images to maximize its function.

As a primary study, this research has a comparatively limited sample size. Future studies should aim to recruit more participants to generate more comprehensive results. Additionally, all the participants in this experiment were Chinese, and the language of the articles was also Chinese. Reading habits may vary among different cultures, and different languages may have varying impacts on reading behaviors. Therefore, future studies could use other languages as stimuli and include participants from different countries to explore potential cultural and linguistic influences on reading comprehension. This would enhance the generalizability of the findings and provide a more comprehensive understanding of the topic.

5. Conclusion

The present research examined how pictures and explanatory texts affect people's reading comprehension. The participants were asked to view a series of stimuli in which some contained only texts, some contained texts and relevant images, and some contained texts and images with explanatory texts. An eye tracker was used to record subjects' gazing behaviors during the process. Quantitative eye-tracking results suggested that the presence of pictures could significantly improve people's reading comprehension and reduce their cognitive load. However, the explanatory texts accompanying the pictures did not further facilitate people's cognitive process.

References

[1] Vansteelandt, I., Keer, H. V., & Mol, S. E. (2022). Pre-service teachers' reader profiles: stability and change throughout teacher education. *Journal of Research in Reading, 45*(1), 1-19.

- [2] Gandelman, C. (1991). *Reading Pictures, Viewing Texts*. Gandelman1991-GANRPV.
- [3] Pressley, G. M. (1976). *Mental imagery helps eight-year-olds remember what they read*. *Journal of Educational Psychology*, 68(3), 355–359.
- [4] Samuels, S.J. (1970). *Effects of Pictures on Learning to Read, Comprehension and Attitudes*. *Review of Educational Research*, 40(3).
- [5] Samuels, S. J. (1967). *Attentional process in reading: The effect of pictures on the acquisition of reading responses*. *Journal of Educational Psychology*, 58(6, Pt.1), 337–342.
- [6] Beymer, D., Orton, P. Z., Russell D. M. (2007) *An Eye Tracking Study of How Pictures Influence Online Reading*. In *Human-Computer Interaction – INTERACT 2007: 11th IFIP TC 13 International Conference, Rio de Janeiro, Brazil, September 10-14, 2007, Proceedings, Part II 11* (pp. 456 460). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-74800-7_41
- [7] Cheng Li & Yang Zhiliang. (2006). *Eye movement studies of college students reading illustrated articles*. *Psychological science* (03), 593-596 + 562. doi:10.16719/j.cnki.1671-6981.2006.03.017
- [8] Liu Baoxia, Zhang Houcan, Shu Hua. (1996). *The role of the illustrations in expository reading comprehension*. *Psychology Journal* (02)
- [9] Long Liqun. (2022). *Applied Research on Illustration Teaching (Master's dissertation, East China Institute of Technology)*. <https://43.138.60.40/KCMS/detail/detail.aspx?dbname=CMFD202301&filename=1022667284.nh>
- [10] Meng Xiaoli. (2018). *Research on illustrations of compulsory Chinese textbooks (master's thesis, Guangxi Normal University)*. <https://43.138.60.40/KCMS/detail/detail.aspx?dbname=CMFD201901&filename=1018851332.nh>
- [11] Wang Fuxing, Duan Zhaohui, Zhou Zongkui & Chen Jun. (2015). *The effect of proximity effects on graphic integration in multimedia learning: the role of cues*. *Psychological Journal* (02), 224-233.