Practical Research on AI Visual Focus Analysis in Online Teaching

Ming-Feng Lee [1], Guey-Shya Chen[2], Ming-Zhi Cheng*[3], Hui-Chien Chen[4], and Jian-Zhi Chen[5]

[1] Department of Information Management, National Taichung University of Science and Technology, Taiwan
antonio6561@gmail.com
[2,3,4] Institute of Educational Information and Statistics, National Taichung University of Education, Taiwan
grace@mail.ntcu.edu.tw
*cms109101@gm.ntcu.edu.tw
gigi.baby@yahoo.com.tw
[5] Corporate Synergy Development Center(CSD), Taiwan
t296210@gmail.com

Abstract. This project aims to study learners' visual focus in online learning and their behavior patterns when engaging with different learning media and cognitive processes. It uses AI models and data analysis methods to evaluate learners' online preferences and information processing.
The project consists of three stages：
(1) Collecting and Integration: Use visual movement devices to collect learners' visual movement data and other data to analyze visitors' visual behavior and identify key eye movement factors.
(2) AI Applications: Use AI data mining technologies to identify the visual focus of the target learners through machine learning and data analysis.
(3) Integrated Analyses of Online Learning: Optimize online learning design by combining visual focus analysis and information collected in visual mode. Collaborate with companies to establish an "adoption model for visual focus analysis of online education," which will be verified by the online teaching platform.

Keywords: Visual focus, Online learning, Artificial intelligence, Data analysis, behavior patterns.

1 Introduction

This project aims to achieve two outcomes：

(1) thorough research on the visual focus process of learners in online learning

(2) creation of an "application model of online teaching visual focus analysis" in collaboration with industry, to enhance the online learning experience for learners.

1. This project analyzes learners' eye movement patterns to categorize their learning situations and develop practical applications. It also aims to improve deep learning and

data mining models through a visual focus analysis system. The project builds upon previous media detection applications and expands its model.

2. This project aims to study the differences between online and physical learning by conducting eye movement observation research on online learning. The research results will be shared through academic speeches, workshops, and industry-government-academia lectures to promote interdisciplinary concepts and the transformation of research innovations into interactive online education displays and technology. The project will also integrate various resources to work together.

3. Developmental psychology studies show that the human brain analyzes gaze to predict mental behavior and manipulate colors to attract consumers. Studies are conducted using questionnaires, stimuli, and research methods. In addition, psychologists have found that visual behaviors may also show a degree of preference, for example: they will look twice as long at favorite items (Li Suxin and Li Jimian, 2001), or their pupils will dilate significantly when they see favorite items (Janisse & Peavler, 1974).

This study aims to use a webcam eye tracker and questionnaire survey to monitor real-time cognitive processes and attention, providing a reference for teachers designing teaching materials in hybrid teaching. The study has two purposes: exploring the influence of different design styles of pictures and texts on subjects' reading and exploring the relationship between graphics and cognition using the eye movement data of subjects.

2 Literature discussion

2.1 A Subsection Sample

Eye Movement Research on Graphic Reading Most studies on graphic reading are based on Mayer's cognitive theory of multimedia learning (CTML) (Mayer, 2005, 2014) and Schnotz's integrated model of graphic comprehension (Schnotz et al., 2014). According to the dual coding theory (Paivio, 1986), the human cognitive system comprises two independent but interrelated subsystems: the language system and the image system. This theory assumes that visual images and Chinese characters require different processing channels. CTML is based on three basic assumptions. First, human information processing involves two independent and symmetrical channels: auditory and image. Second, memory capacity is limited, with individuals able to remember about 7±2 elements in each system, according to Miller's (1956) and Simon's (1974) Memory Span Test. Third, learners actively construct knowledge by selecting, organizing, and integrating information and prior knowledge to create a coherent mental model. Empirical research supports CTML's assumption that combining pictures and text enhances learning compared to text-only presentation. Mayer suggests that images facilitate high-level learning by helping readers form a mental model of concepts. Since the cognitive system's working memory has limited capacity for retention and operation, presenting visual and textual information together allows for processing and encoding of information through the dual channels of spoken language and images. This approach reduces cognitive load and promotes learning (Wang Zining & Jian Yuqin, 2022) .

Current literature on eye movement research in graphic reading can be broadly categorized into two directions. The first direction, business psychology, explores whether people follow specific eye movement patterns when viewing advertisements or browsing the Web, and which objects attract their attention. The second direction, cognitive learning, examines the cognitive processes involved in scientific graphic reading, including the reader's reading path and how information is processed. It also investigates how graphics aid in the integration of graphics and text, such as whether pictures enhance understanding of concepts and whether important information marked on pictures aids in comprehension.

2.2 Eye Tracker

Eyes are crucial for communication with the world, and eye tracking technology can measure gaze points (Cornsweet, 1958). Eye tracking research on reading behavior focuses on fixations, saccades, and regressions (Poole & Ball, 2005). The number of gaze times in a specific area indicates its importance, with longer gaze times suggesting more challenging information extraction or more attractive targets. For informational pictures, more gaze times may indicate complexity and multiple processing cycles. Time spent gazing correlates with gaze length during the gaze period (Just & Carpenter, 1976).

Hyona, Lorch, and Kaakinen (2002) found that personal reading strategies, knowledge, and experience influence gaze time and position when viewing pictures or text. Chen et al. (2014) used eye-tracking to investigate the impact of different presentation modes on college students' performance on computerized assessments, and found that rereading time predicted their evaluation performance. Students with longer gaze times and shorter saccade distances performed better on physical concept tasks, indicating that pictures convey physical concepts faster and more effectively than text. Eye-tracking can provide substantial evidence for understanding students' performance.

3.Investigating the Reading Process using Eye-Tracking Technology. Eye-tracking technology is used to track the movement of the eyeballs and link them with stimuli, allowing researchers to determine where individuals place visual attention on the stimulus, how long they sustain their focus, and the sequence of their gaze patterns (Holmqvist et al., 2011). Metrics include fixation, saccade, and gaze. Gaze refers to almost stationary eyes that retrieve information from the stimulus. Position and duration of gaze indicate which information the individual has accessed and how deeply they have processed it. The fixation point duration usually ranges from 100 to 500 milliseconds (Rayner, 1998). Nearly 100 indicators, such as AOI and POI, including observation time, are used for detailed analysis of eyeball trajectories (Chen Xuezhi et al., 2010).

It is suggested that the use of graphic and text materials in learning has a better effect than using only text materials (Zhang & Gu, 2011). Eye-tracking metrics have been used to explore the impact of news and text layout on attention distribution, and changes in the relative positions of pictures and text have been found to affect attention distribution (Tang & Zhuang, 2005). Design principles of teaching materials have been classified based on six eye movement studies related to graphics and text (Mayer, 2010).

Eye-tracking technology has been used to record students' eye movement behavior and provide important indicators such as gaze time, back gaze, and gaze shift, which reflect attention distribution and cognitive processes (Henderson & Hollingworth, 1999; Josephson & Holmes, 2002; Rayner & Pollatsek, 1987; Yang & McConkie, 1999).

3 Research Methods

3.1 1.Research Participants

This eye movement experiment was conducted during data mining. The research participants were self-selected students, with an effective sample of 44 participants, including 16 females and 28 males. Each subject completed the eye movement experiment and then answered a viewing questionnaire.

3.2 Research Tools

Eye-tracking measurement instrument

This study utilized Real-eye Web eye tracking software, where participants only needed a network camera to collect data. The software performed eyeball correction before data collection. The data collected included hotspots and eye movement tracks, which recorded the time and frequency of participants' gaze and scanning behavior.

3.3 Defining the area of interest (AOI) in images.

Research materials were taken from online coffee resources, and subjects were placed in a cafe environment through text descriptions. Different environmental elements and coffee cup colors were used to assess how they affect mood and coffee perception. Figures 1 and 2 presented two different atmospheres, cool and warm, and identified areas of interest including coffee, coffee pots, tablets, books, and coffee cups of different sizes, blue cups, water cups, and desserts.

Fig. 1. Display of coffee book and tablet in black and white color style.

Fig. 2. Display coffee, water glasses and bread in a wood tone color scheme.

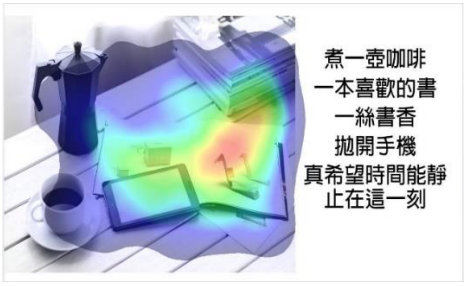

Fig. 3. The heat map of eye movement analysis in Figure 1..

Fig. 4. The heat map of eye movement analysis in Figure 2.

4 Discussion and suggestion

This study uses eye-tracking to study how graphic designs affect cognition, emotion, and perception. Colors have strong emotional impacts, such as red for excitement, orange for stimulation, and blue for comfort and safety (Ballast, 2002; Wexner, 1982). Studies show that people of all ages, races, and education levels prefer culturally associated colors (Adams & Osgood, 1973; Eysenck, 1941).

The study found that subjects spend less time looking at cool-colored designs, feeling calm and relaxed, and perceiving coffee as bitter. On the other hand, warm-colored designs lead to a happy and relaxed emotional state and a perception of coffee as sweet. The hot zones in the coffee cups were also more appealing to the participants. These

findings can be used as a reference for designing teaching materials with consideration of color configuration to enhance learners' interest.

References

1. Jian, YC, & Wu, C.-J. (2012). The effect of arrows in an illustration when reading scientific text: Evidence from eye movements and reading tests. Chinese Journal of Psychology, 54, 385–402. https://doi.org/10.6129/cjp.2012.5403.07
2. Adams, F. M., & Osgood, C. E. (1973). A cross-cultural study of the affective meanings of color. Journal of Cross-Cultural Psychology, 4(2), 135-156.
3. Ballast, D. (2002). Interior design reference manual. Professional Pub. Inc.: Belmont, CA.
4. Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/ correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Erlbaum.
5. Eysenck, H.J. (1941). A critical and experimental study of colour preferences. The American Journal of Psychology, 385-394.
6. Guo, D., Zhang, S., Wright, K. L., & McTigue, E. M. (2020). Do you get the picture? A meta-analysis of the effect of graphics on reading comprehension. AERA Open, 6(1). Advance online publication. https:// doi.org/10.1177/2332858420901696.
7. Hegarty, M., Canham, M. S., & Fabrikant, S.I. (2010). Thinking about the weather: How salience and knowledge affect performance in a graphic inference task. Journal of Experimental Psychology: Learning, Memory and Cognition, 36(1), 37-53. doi:10.1037/ a0017683
8. Janisse, M. P., & Peavler, W. S. (1974). Pupillary Research Today: emotion in the eye,Psychology Today, 7, 60-63.
9. Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., et al. (2013). A re- view of using eye-tracking technology in exploring learning from 2000 to 2012. Educational Research Review, 10, 90-115. doi:10.1016/j.edurev.2013.10.001
10. McTigue, E. M. (2009). Does multimedia learning theory extend to middle-school students? Contemporary Educational Psychology, 34, 143-153.https://doi.org/10.1016/j.cedpsych.2008.12.003
11. Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), The Cambridge Handbook of multimedia learning (pp. 31–48). Cambridge University Press.
12. Mayer, R. E. (2014). The Cambridge handbook of multimedia learning (2nd ed.). Cambridge University Press.
13. Paivio, A. (1986). Mental representations: A dual-coding approach. Oxford University Press.
14. Segers, E., Verhoeven, L., & Hulstijn-Hendrikse, N. (2008). Cognitive processes in children's multimedia text learning. Applied Cognitive Psychology, 22, 3375-387. https://doi.org/10.1002/acp.1413
15. Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), The Cambridge Handbook of multimedia learning (2nd ed., pp. 72–103). Cambridge University Press.
16. Wexner, L. B. (1982). The degree to which colors (hues) are associated with mood-tones. Journal of applied psychology, 38(6), 432.