

# The dilation of the pupils and the mystery of gaze: Revealing the microcosm of cognitive load

Caijilahu Bao\*, Lulu Zhang

Inner Mongolia University of Technology, Department of Data Science and Applications, Hohhot, China, 0471-3602107

\* Corresponding Author: caijilahubao@imut.edu.cn

## ABSTRACT

In the process of education and learning, the lack the process of education and learning, the lackers' cognitive load may lead to the lack of accurate assessment of learners' cognitive load may lead to the design of teaching content, the provision of learning resources, and the choice of teaching methods that do not match the cognitive characteristics of learners, thereby causing difficulties in learning and affecting learning outcomes. To better achieve teaching objectives, this article reveals the level of cognitive load faced by participants in searching for specific texts or symbols through tracking and in-depth analysis of eye parameters such as pupil diameter, gaze point, and gaze duration. The results show that when individuals face complex tasks, their pupils dilate sharply, a phenomenon stemming from the demand for more cognitive resources by complex tasks, which triggers adjustments in the autonomic nervous system. In addition, high-frequency scanning behavior was also observed, and this frequent scanning behavior is precisely the user's attempt to quickly capture and process a large amount of information in a limited time, reflecting the tense allocation and efficient utilization of cognitive resources. The effective combination of the two can better assess the cognitive load level of the subjects.

**Keywords:** cognitive load, eye-tracking, pupil diameter, fixation path, saccade.

## 1. INTRODUCTION

In the context of the information age, the assessment of cognitive load in various task scenarios has become an important research topic in the fields of psychology, education, and human-computer interaction. Cognitive Load Theory (CLT)[1] refers to the amount of mental resources required for an individual to process information and engage in thinking. This concept was introduced by educational psychologist John Sweller in 1988, primarily to describe the cognitive demands during the learning process. Therefore, accurately measuring and assessing cognitive load is crucial for optimizing task design, enhancing learning outcomes, and improving human-computer interaction experiences.

Eye-tracking technology, as a non-invasive measurement tool, has been widely used in cognitive load research in recent years. By recording eye movement data of individuals during task execution, it is possible to infer their attention allocation and information processing, thereby indirectly assessing their cognitive load levels. Among various eye-tracking metrics, pupil diameter and fixation point are considered two important indicators for evaluating cognitive load[2].

Changes in pupil diameter are often closely related to an individual's cognitive load. Fixation points reflect the individual's attention allocation and information processing methods during task execution. By analyzing the distribution and duration of fixation points, we can understand the degree of attention concentration and information processing strategies of the individual, thereby inferring their cognitive load levels. Combining these two indicators allows for a more accurate estimation of cognitive load through eye-tracking.

In this study, we delve into the cognitive load levels faced by participants during tasks involving the search for specific words or symbols.

## 2. RELATED WORK

Divya Venkatesh et al.[3] ingeniously utilized eye-tracking technology to reveal cognitive load levels in manual text classification tasks, as evidenced by recorded fixation counts, fixation durations, and pupil diameter changes. They found

that native English speakers exhibited lower cognitive load when performing simple tasks, but experienced a significant increase in cognitive load when handling complex, more challenging narrative texts. Saiz-Manzanares et al.[4] extensively explored students' learning performance in virtual laboratories and its impact on learning outcomes using eye-tracking technology. The study also analyzed the role of prior knowledge in learning outcomes. Researchers identified clusters closely associated with cognitive load in both related and unrelated interest areas. Results indicated significant improvements in learning outcomes for students engaged in virtual laboratories, suggesting the effectiveness of this learning approach. Notably, prior knowledge did not significantly affect students' cognitive load during the learning process. Additionally, the study revealed multiple clusters associated with different interest domains, closely linked to cognitive load indicators in related and unrelated areas.

Lee et al. [5] study revealed, as expected, significant pupil responses to difficult tasks compared to simple tasks. Specifically, pupil responses showed a positive correlation with performance measures only when participants faced difficult tasks. This suggests that pupil response, after controlling for light reflex interference, can serve as an effective measure of cognitive load in virtual reality (VR) training assessments. This finding not only provides empirical support for the relationship between task difficulty and cognitive load but also offers critical insights for optimizing VR training experiences and assessing learning outcomes.

Mishra et al.[6] proposed and evaluated cognitive load modeling methods related to text comprehension by studying the complexity of gaze scan paths during reading. They aimed to quantify cognitive effort in the reading process by modeling readers' eye movement behaviors. Results indicated that measuring scan path complexity using eye-tracking records could generate better cognitive models and explain users' reading effectiveness.

Majoomi et al.[7] investigated the impact of layout on audience comprehension and cognitive load in information graphics. They analyzed eye-tracking data and provided quantitative evidence of how layout variations affect participants' understanding and cognitive load. The results showed that jagged layouts led to higher comprehension and lower cognitive load.

### 3. METHODOLOGY

#### 3.1 Participants and Experimental Apparatus

The participants were 15 normal-sighted university students and community members who volunteered for the experiment. Due to technical issues (mainly eye tracker calibration) or misunderstandings about the tasks, data from 3 participants were discarded. The final sample comprised N=12 participants, aged between 20 and 50 years (M=28.5; SD=7.98). Eye movements were recorded binocularly using an iView X Hi-Speed (SMI) eye tracker. During the recording, the head of each participant was fixed by a chin rest (Figure 1).

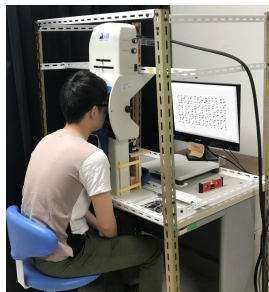


Figure 1. The participant during the experimental procedure.

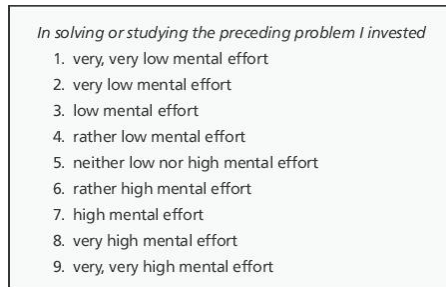


Figure 2. The Paas (1992) subjective rating scale.

#### 3.2 Scale

The PAAS Self-Assessment Multidimensional Scale[8], proposed by Paas in 1992, is a subjective evaluation method for measuring cognitive load. The PAAS uses a 9-point Likert scale as shown in Fig 2, ranging from very very low (1) to very very high (9), to assess the mental effort invested during learning and testing. The "intensity of effort" evaluated subjectively is considered an index of cognitive load. Some studies use task difficulty as the standard for assessing cognitive load. Whether using mental effort or difficulty level, subjective evaluation scales can sensitively reflect differences in cognitive load. The PAAS scale has been proven to have high reliability. Compared to physiological

measurements, the PAAS self-assessment scale is more sensitive to the detection of cognitive load and has less interference.

## 4. EXPERIMENT

The working memory capacity (WMC)[9] of each participant, as a controlled independent variable, was measured through a reading span test[10] before the main experimental task. The reading span test required participants to read a series of sentences and remember the last word of each sentence. After completing one set of sentence-final word recall, another set of sentences would be presented for reading. The WMC score was calculated as the total number of correctly recalled words throughout the entire test.

### 4.1 Initial pupil size measurement and baseline data collection

Before the formal experiment, the initial pupil size of each participant was measured. Participants were required to focus on a fixed point in the center of the screen while maintaining stability of the eyes and head. The measurement time was approximately 15 seconds, and the original pupil diameter was recorded to obtain baseline data, ensuring that changes in pupil size were caused by the task.

### 4.2 Experimental procedure

Before the experiment begins, the experimenter will explain the details of the experimental procedure and precautions to all participants, including the purpose of the experiment, task requirements, maintaining a stable posture, concentrating attention, and suggestions for relaxation during breaks. Participants are required to sign a consent form, provide basic information, and complete the reading span test.

The reading span test consists of test materials with sets of 2 to 6 sentences, with 5 sets of sentences per group. After reading each set, participants need to judge whether a sentence's meaning is consistent with the sentences they have read and recall the last word of each sentence in sequence.

Upon completing the WMC, participants began the main task with their eyes tracked after a 5-point calibration. The main experiment was divided into six groups, with a rest period between the third and fourth experiments, and eye calibration conducted before the first and fourth experiments. After calibration, participants saw a target character or symbol at the center of the screen, which they were required to remember. This target character could be from a text dial. At the start of the search task, the participant's eye search trajectory was recorded on the screen displaying the text dial. Participants had 10 seconds to memorize the target. Subsequently, a dense image containing multiple distractor images appeared on the screen, and participants had 60 seconds to locate the target character or symbol.

The text dial was displayed on a 27-inch monitor (1920 pixels wide  $\times$  1080 pixels tall, or 524 mm wide  $\times$  295 mm tall), with room illumination of about 270 lx. Participants viewed the center of the monitor as the initial state, with a distance of approximately 600 mm between them and the monitor. Characters were arranged in a 16 $\times$ 16 square area, with each character being 8 mm square on the monitor, and the distance between characters also being 8 mm. A total of 256 characters were selected for random arrangement from among Hiragana, Katakana, Roman letters, Greek letters, Cyrillic letters, Kannada letters, and other symbols. Similar-looking characters such as the digit "0" and the letter "O" were excluded. To disrupt the arrangement, the position of each character was randomly reconfigured up to 50% from the center position in all directions. This was done to prevent gaze points from being induced toward specific directions like horizontal or vertical. The experimental equipment continuously recorded changes in pupil size and gaze point data, while also recording reaction time and accuracy rate.

After each experimental session, participants estimated their level of cognitive demand through the PAAS questionnaire. The analysis of pupil diameter changes during task execution, as well as gaze point distribution, dwell time, and scan paths, was used to evaluate attentional focus and information processing strategies. The completion time for each task round and the accuracy rate of target recognition were also recorded.

## 5. DATA ANALYSIS

After each group of experiments, collect the PAAS scale data filled out by the participants. All data will be collected and organized for analyzing the relationship between pupil diameter and gaze points with cognitive load. The subjective cognitive load assessment results will be compared with the objective measurement data to analyze the correlation between the two.

### 5.1 The change in pupil size

Firstly, in the preprocessing step, eye-tracking data is extracted to remove data 200 milliseconds before and after the beginning of blinks identified by the eye tracker. After this preprocessing step, we calculate the inter-experimental changes in pupil diameter related to the baseline experiment. Before calculating metrics based on changes in pupil diameter, a Butterworth smoothing filter is applied to the raw pupil diameter data. The parameters of the Butterworth filter are chosen to remove high-frequency noise observed in the signal. We take the input signal  $x(t)$  and produce its filtered version,  $\hat{x}(t)$ , as the output, where  $\hat{\phantom{x}}$  denotes smoothing. We use a  $2^{nd}$  degree Butterworth filter with the critical frequency set to  $1/4$  of a half-cycle per sample, which is  $1/8^{th}$  of the sampling period (the point where the gain drops to the passband  $1=2$ ). That is to say, representing the pupil diameter signal as  $x(t)$ , the signal is smoothed (to order  $s$ ) by convolving  $2p+1$  inputs of  $x_i$  with the filter  $h_i^{r,s}$  and  $2q+1$  (previous) outputs of  $\hat{x}_i$  with the filter  $g_i^{r,s}$  at midpoint  $i$ :

$$\hat{x}_n^s(t) = \frac{1}{(\Delta t)^s} \left( \sum_{i=-p}^p h_i^{r,s} x_{n-i} - \sum_{i=-q}^q g_i^{r,s} \hat{x}_{n-i} \right) \tag{1}$$

Where  $r$  and  $s$  denote the order of polynomial fitting to the data and its derivatives.

### 5.2 Measurement of eye gaze point position and length of gaze point movement.

The exploration time from the eye tracker is used as the eye movement data, and the pupil diameter at discrete time  $t$  and the gaze point coordinates  $(u_t, v_t)$  on the display are converted into pixels as shown in Fig 3. To obtain gaze point coordinates  $(x_t, y_t)$  in millimeters, the display size is standardized as 1,920 pixels wide (524 mm) and 1,080 pixels tall (295 mm). The conversion method is calculated using the following formulas (Formulas 2, 3).

$$x_t = 524u_t / 1920 \tag{2}$$

$$y_t = 295v_t / 1080 \tag{3}$$

One pixel corresponds to approximately 0.03 degrees of the eye angle, and similarly, 1 mm is equivalent to approximately 0.1 degrees. The length  $r_t$  of the gaze point movement between discrete time points  $t-1$  and  $t$  is calculated using the formula (equation 4) based on the coordinates  $(x_t, y_t)$  at time  $t-1$  and  $(x_{t-1}, y_{t-1})$  at time  $t$ .

$$r_t = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2} \tag{4}$$

When the pupil diameter is measured as 0, it means that the experimental participant involuntarily blinked. Such gaze point coordinates are excluded from the analysis.

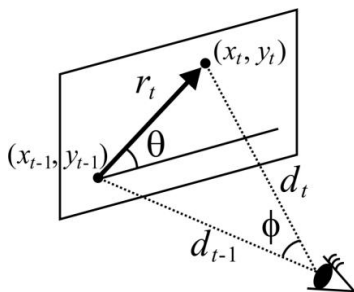


Figure 3. Schematic diagram of eye angle and saccade angle  $\theta$ .

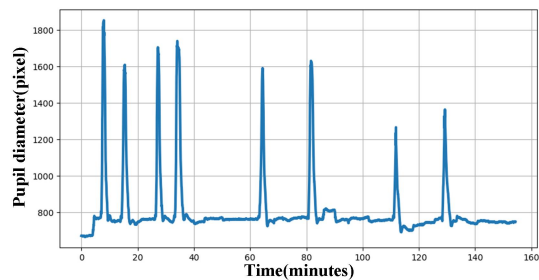


Figure 4. Complete Pupil Size Graph for the Experiment.

## 6. SUMMARY AND DISCUSSION

### 6.1 Pupil diameter change

In this experiment, we recorded the pupil size changes of the subjects during the process of memorizing and searching for target words using an eye tracker. The experiment was divided into four key stages: the preparation stage, the memory stage, the search stage, and the instant of finding the target. Fig 4 depicts the pupil change graph during the experiment.

During the preparation stage, when the subjects were gazing at the screen and preparing to receive the task, the pupil size remained at a baseline state with no significant changes. This stage was intended to allow the subjects to calm down.

During the memory stage, the subjects focused their attention on memorizing the target words or symbols. Due to the increase in cognitive load, the pupil size often expanded, reflecting an increase in the brain's resources required for information processing.

In the search stage, the screen displayed a dial containing distractor images or text, and the subjects needed to find the target word. As the task difficulty and distractions increased, the subjects' cognitive load further escalated, leading to a significant dilation of the pupils. When the task was more difficult, the pupil dilation was more pronounced (as shown in Fig 5a); while for easier tasks, the pupil changes were smaller (as shown in Fig 5b).

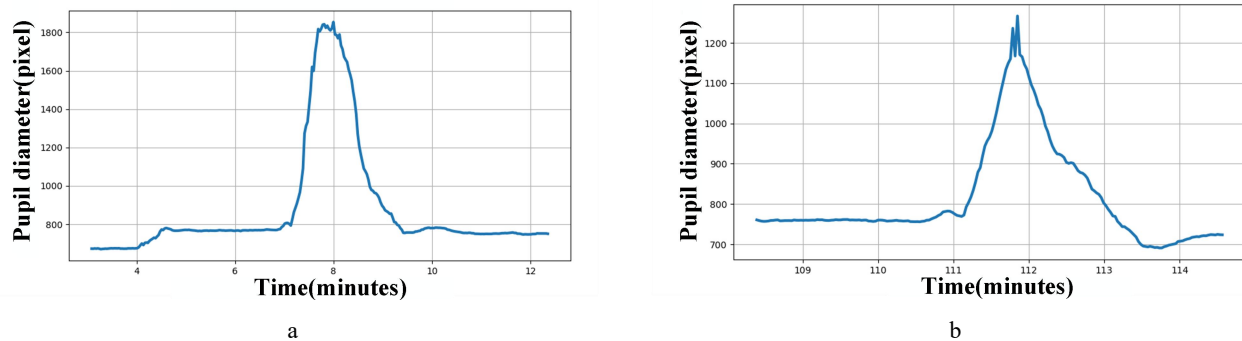


Figure 5. Partial Pupil Diameter Change Graph.(a.Significant pupil dilation(increasing to 1800).b.Relatively insignificant pupil dilation(increasing to 1300).)

Finally, at the instant when the subject found the target, there was a significant change in pupil size. Typically, after finding the target, the cognitive load rapidly decreased, and the subject's pupils gradually returned to a more relaxed state. The pupil changes during this stage reflect the subject's cognitive state adjustment after completing the task.

By observing and analyzing the pupil size changes during these stages, we can gain a deeper understanding of the subjects' cognitive load and attention states under different task scenarios. The changes in pupil size not only reveal the impact of task difficulty on the subjects but also provide insights into their psychological states and response patterns in complex cognitive tasks.

### 6.2 Changes in the Path of Fixation Points

When users repeatedly fixate on a certain area and revisit it, it often indicates that they are engaging in deep and complex processing or thinking about the information in that area. The path of fixation points jumping from one location to another not only maps out the user's path of information acquisition but also reveals their unique processing strategies. Notably, when users make high-frequency saccades, this is often a significant signal of higher cognitive load as they process complex information (Fig 6).

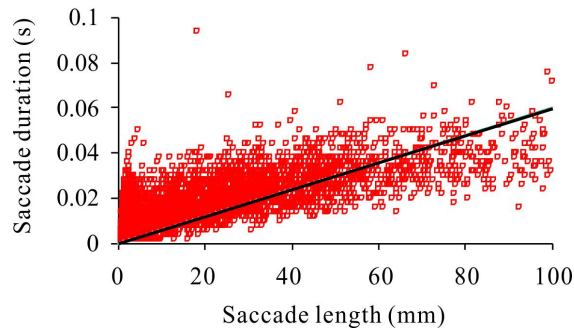


Figure 6. shows the relationship between scan length and scan duration.

As shown in Fig 7, which includes the movement of dials and gaze points, with the length of line segments representing the length of saccades, and the radius of circles indicating the duration of fixational micro-movements. when the subjects successfully found the target characters (Fig 7a), their gaze point images exhibited a clear and focused trend, almost no longer searching downwards; however, when the subjects failed to quickly find the target characters (Fig 7b), their gaze point images appeared relatively chaotic, accompanied by an increase in pupil diameter, which intuitively reflects that their cognitive load was significantly increasing at the moment. This phenomenon not only reveals the psychological state of users when processing information but also provides valuable clues for researchers to gain a deeper understanding of user behavior and psychological activities.

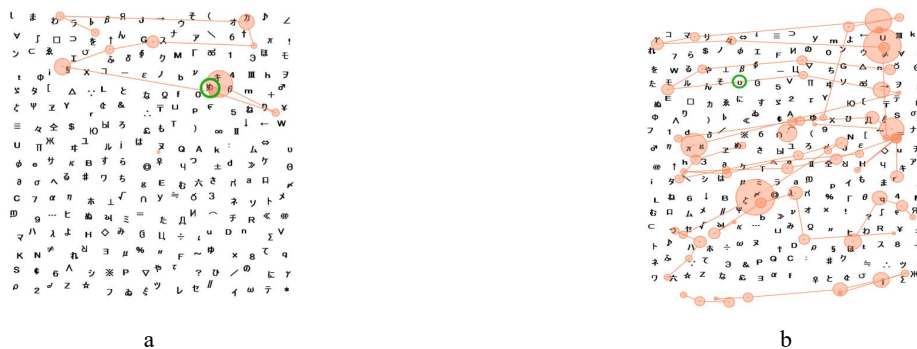


Figure 7. An example of the validation dataset of the present invention. (a.Validation dataset under low cognitive load; b.Validation dataset under high cognitive load.)

As shown in Fig 8、9, the frequency and angular distribution of saccade lengths are presented. In Fig 8, the frequency of saccades with a length of 2 mm is 482 (only three frequency numbers are displayed in the figure), and the maximum saccade length is 94 mm. Fig 9 displays the frequency of saccade angles every 10 degrees, such as the frequency of saccades at an angle of 10 degrees being 33. Fig 10 represents the distribution of fixational micro-movement durations. The figure shows that the minimum duration of fixational micro-movements is 0.002 seconds and the maximum is 0.334 seconds. For example, the frequency of fixational micro-movements with a duration of 0.002 seconds is 134 (only three frequency numbers are indicated in the figure).

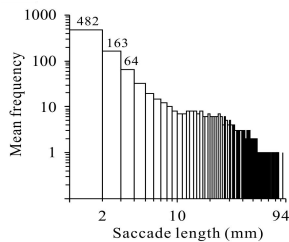


Figure 8. Distribution of the average saccade length.

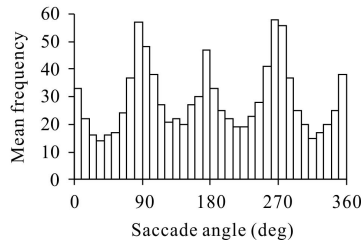


Figure 9. Distribution of the average saccade angle.

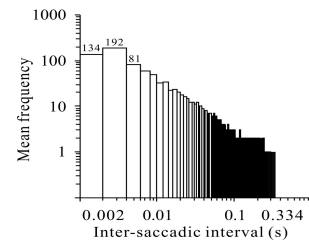


Figure 10. shows the distribution of fixational micro-movement durations.

## 7. CONCLUSION

Pupil diameter and fixation points in eye movement analysis serve as key indicators for assessing cognitive load, which is crucial to understanding users' cognitive processing. By processing and analyzing users' eye movement data and their completed PAAS scales, we can demonstrate that pupil diameter and fixation points in eye movement data can be used to detect cognitive load during complex or simple tasks. Research results indicate that repeated fixations on specific areas reflect in-depth information processing and complex thinking, revealing the focus and depth of information processing. Fixation jumps reflect users' switching strategies between different information sources to form a more comprehensive cognitive picture. High-frequency scanning behavior indicates a high cognitive load faced by users when processing complex information, while pupil dilation provides an objective physiological basis for assessing this load. Combining pupil dilation, fixation duration, and subjective feedback allows for a comprehensive assessment of users' cognitive load status. Therefore, eye movement analysis based on pupil diameter and fixation points offers an important perspective for understanding users' cognitive processing, contributing to the precise evaluation of cognitive load and information processing strategies.

## REFERENCES

- [1] J. J. C. s. Sweller, "Cognitive load during problem solving: Effects on learning," *Cognitive science*, vol. 12, no. 2, pp. 257-285, 1988, doi: [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7).
- [2] L. Ziaka, A. J. J. o. E. P. H. P. Protopapas, and Performance, "Cognitive control beyond single-item tasks: Insights from pupillometry, gaze, and behavioral measures," *Journal of Experimental Psychology: Human Perception*, vol. 49, no. 7, p. 968, 2023, doi: [10.1037/xhp0001127](https://doi.org/10.1037/xhp0001127).
- [3] J. Divya Venkatesh, A. Jaiswal, and G. Nanda, "Analyzing Cognitive Load Associated with Manual Text Classification Task Using Eye Tracking," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2023*, vol. 67, no. 1, pp. 193-198: SAGE Publications Sage CA: Los Angeles, CA, doi: [10.1177/21695067231192221](https://doi.org/10.1177/21695067231192221).
- [4] M. C. Sáiz-Manzanares, R. Marticorena-Sánchez, L. J. Martin Anton, I. González-Díez, and M. Á. J. I. J. o. H. C. I. Carbonero Martín, "Using eye tracking technology to analyse cognitive load in multichannel activities in University Students," *International Journal of Human-Computer Interaction*, pp. 1-19, 2023, doi: [10.1080/10447318.2023.2188532](https://doi.org/10.1080/10447318.2023.2188532).
- [5] J. Y. Lee, N. de Jong, J. Donkers, H. Jarodzka, and J. J. J. I. T. o. L. T. van Merriënboer, "Measuring Cognitive Load in Virtual Reality Training via Pupillometry," *IEEE Transactions on Learning Technologies*, 2023, doi: [10.1109/tlt.2023.3326473](https://doi.org/10.1109/tlt.2023.3326473).
- [6] S. Mathias, D. Kanojia, K. Patel, S. Agarwal, A. Mishra, and P. J. a. p. a. Bhattacharyya, "Eyes are the windows to the soul: Predicting the rating of text quality using gaze behaviour," *arXiv preprint arXiv:1810.04839*, 2018, doi: [arxiv-1810.04839](https://doi.org/10.1810.04839).
- [7] A. Majooni, M. Masood, and A. J. I. V. Akhavan, "An eye-tracking study on the effect of infographic structures on viewer's comprehension and cognitive load," *Information Visualization*, vol. 17, no. 3, pp. 257-266, 2018, doi: [10.1177/1473871617701971](https://doi.org/10.1177/1473871617701971).
- [8] F. Paas, P. Ayres, and M. J. R. I. i. E. T. T. F. S. L. Pachman, Information Age Publishing Inc., Charlotte, NC, "Assessment of cognitive load in multimedia learning," *Recent Innovations in Educational Technology That Facilitate Student Learning*, Information Age Publishing Inc., Charlotte, NC, pp. 11-35, 2008.
- [9] E. Navarro, H. Hao, K. P. Rosales, and A. R. J. B. R. M. Conway, "An item response theory approach to the measurement of working memory capacity," *Behavior Research Methods*, vol. 56, no. 3, pp. 1697-1714, 2024, doi: [10.3758/s13428-023-02115-3](https://doi.org/10.3758/s13428-023-02115-3).
- [10] S. Wang, L. L. Wong, and Y. J. I. J. o. A. Chen, "Development of the mandarin reading span test and confirmation of its relationship with speech perception in noise," *International Journal of Audiology*, pp. 1-10, 2024, doi: [10.1080/14992027.2024.2305685](https://doi.org/10.1080/14992027.2024.2305685).