An Eye-Tracking-Based Evaluation Model for Information Visualization Design Effectiveness

Na Wang, Weiqian Mao, Shengnan Wu*, Jiahui Wang

Harbin Institute of Information Technology, Harbin, China

*945327072@qq.com

Abstract. This study proposes an eye-tracking-based model to assess information visualization effectiveness. By integrating eye movement metrics—First Fixation Duration (FFD), Total Fixation Duration (TFD), Scanpath Length, and Entropy—with task performance (Accuracy Rate), a composite index termed Visualization Effectiveness Index (VEI) is constructed. A controlled experiment with expert and novice users across multiple visualization types (tree diagrams, heatmaps, and Sankey diagrams) shows that the model effectively differentiates design approaches and provides insights into visual efficiency, cognitive load, and user comprehension. Compared with traditional subjective methods, it improves reliability and reproducibility and supports data-driven visualization design optimization.

Keywords: Eye-tracking; Information visualization; Design evaluation; Visual attention; Visualization effectiveness index.

1 Introduction

Information visualization is widely used in fields such as finance, healthcare, and public communication to convey complex data efficiently[1]. However, its effectiveness is still mainly evaluated through subjective methods, such as expert reviews and user surveys, which lack objective and quantifiable evidence[2]. This limitation hinders systematic understanding of how visual design influences user cognition and information processing. To address this issue, eye-tracking technology captures users' visual attention and cognitive behaviors in real time[3]. Therefore, this study proposes an eye-tracking-based evaluation model that integrates eye movement metrics with task performance data to provide an objective framework for assessing and optimizing visualization design effectiveness[4].

R. Zur et al. (eds.), *Proceedings of the 2026 5th International Conference on Art Design and Digital Technology (ADDT 2026)*, Advances in Computer Science Research 131,
https://doi.org/10.2991/978-94-6239-737-8_18

2 Literature Review

2.1 Eye-Tracking in Visualization and Design Evaluation

Eye-tracking technology has been widely applied in cognitive psychology and human–computer interaction to investigate visual attention and information processing. The “eye–mind” hypothesis proposed by Just and Carpenter suggests a strong correlation between fixation behavior and cognitive processing, providing a theoretical foundation for using eye movement data to evaluate user interaction with visual content[5]. Recent studies have extended this approach to interface and visualization design, demonstrating that metrics such as fixation duration, scanpath, and gaze distribution can effectively reflect users’ perceptual and cognitive responses.

In the field of information visualization, researchers have employed eye-tracking to analyze how users interpret charts, diagrams, and dashboards. Findings indicate that users’ gaze patterns often deviate from designers’ intended visual hierarchy, revealing potential issues in layout and emphasis. However, most existing studies focus on descriptive analysis rather than constructing systematic evaluation models[6].

2.2 Evaluation Methods for Information Visualization

Traditional evaluation methods for information visualization primarily rely on expert reviews, usability testing, and questionnaire-based feedback. While these approaches provide valuable qualitative insights, they are inherently subjective and limited in capturing real-time cognitive processes. Some recent studies have attempted to incorporate quantitative metrics, such as task accuracy and response time, but these indicators alone are insufficient to fully explain user perception and attention mechanisms[7].

Moreover, existing evaluation frameworks often lack integration between physiological data and performance outcomes, and many are developed based on Western user samples, raising concerns about their generalizability. To address these gaps, this study proposes an eye-tracking-based evaluation model that integrates multiple gaze metrics with user performance data. The innovation lies in constructing a quantitative and structured framework that links visual attention patterns to design effectiveness, providing both theoretical and practical contributions to information visualization research.

3 Methodology

3.1 Experimental Design and Data Acquisition

This study proposes a quantitative evaluation framework for information visualization design based on eye-tracking data, as illustrated in Fig. 1. The overall workflow includes visualization input, user interaction, eye-tracking acquisition, data processing, feature extraction, and evaluation model construction.

A total of 60 participants were recruited and divided into two groups: expert users (with design training) and novice users (without professional background). Each participant completed a series of task-oriented experiments using three types of visualizations, including tree diagrams, heatmaps, and Sankey diagrams. The tasks required users to identify key information, compare values, and interpret data relationships, ensuring coverage of different cognitive processes.

Eye movement data were collected using Tobii Pro Glasses 3 at a sampling rate of 100 Hz. A standard calibration procedure was performed before the experiment to ensure accuracy. During the experiment, gaze points, fixation events, and scanpaths were recorded in real time. In parallel, task performance data, specifically accuracy rate, were collected to evaluate user comprehension.

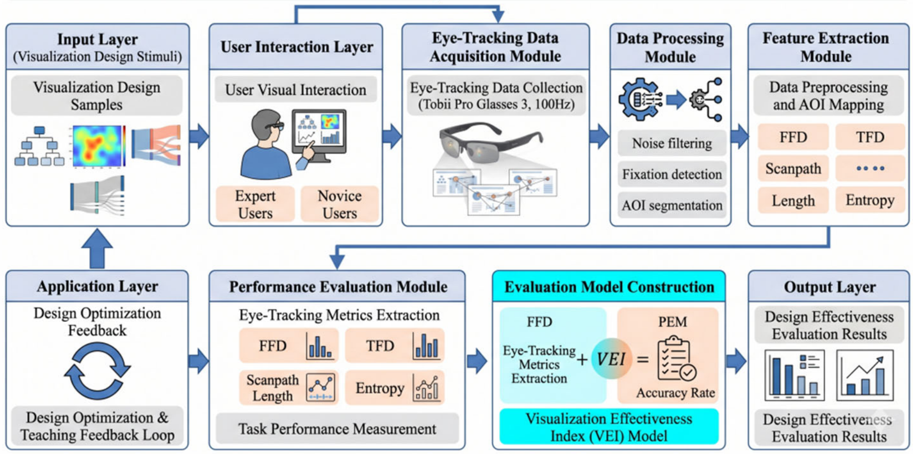

Fig. 1. Eye-Tracking Visualization Evaluation Framework.

3.2 Data Processing and AOI Mapping

The collected eye-tracking data were preprocessed to improve reliability and consistency. A velocity-threshold identification (I-VT) algorithm was applied to classify fixation and saccade events. Fixations were defined as low-velocity gaze points within a minimum duration threshold, while rapid movements were identified as saccades.

To analyze visual attention, Areas of Interest (AOIs) were defined based on key information regions in each visualization. These regions were manually annotated according to design intent, such as titles, legends, and critical data points. The gaze data were then mapped onto the AOIs to determine the distribution of user attention.

After preprocessing, the data were structured into sequential records, including fixation duration, gaze coordinates, and transitions between AOIs. Noise filtering and interpolation methods were applied to remove invalid samples and ensure data completeness. This structured dataset provides a reliable basis for feature extraction and subsequent analysis.

3.3 Eye-Tracking Metrics Extraction

To quantitatively describe user visual behavior, four representative eye-tracking metrics were extracted. First Fixation Duration (FFD) reflects how quickly users attend to key information, indicating the effectiveness of visual guidance. Total Fixation Duration (TFD) represents the overall time spent processing the visualization, which is associated with cognitive load.

Scanpath Length (SL) is used to measure the total distance of eye movement, reflecting the efficiency of visual exploration. A shorter scanpath generally indicates a more coherent and well-structured design. In addition, Entropy (H) is introduced to quantify the distribution of visual attention across different AOIs:

$$H=-\sum_{j=1}^{M} p_j \log p_j \tag{1}$$

where p_j represents the proportion of fixations within the j-th AOI. A higher entropy value indicates a more distributed and balanced allocation of visual attention across different AOIs. This distribution may reflect more comprehensive visual exploration, although its effectiveness depends on task characteristics and does not necessarily imply better performance in all cases.

To evaluate task performance, Accuracy Rate (AR) is used as an objective measure of user understanding. Together, these metrics capture multiple aspects of user interaction, including visual efficiency, cognitive effort, and comprehension outcomes.

3.4 Evaluation Model Construction

To integrate multiple metrics into a unified evaluation framework, this study constructs a composite index termed Visualization Effectiveness Index (VEI). Before aggregation, all metrics are normalized using a min–max method to eliminate scale differences. Negative indicators (FFD, TFD, SL) are inversely transformed, while positive indicators (Entropy and Accuracy Rate) are preserved.

The VEI is defined as:

$$VEI=w_1(1\text{-}FFD')+w_2(1\text{-}TFD')+w_3(1\text{-}SL')+w_4H'+w_5AR' \tag{2}$$

where w_i represents the weight of each metric and $\sum w_i=1$. Equal weights are adopted due to the lack of prior empirical evidence on the relative importance of each metric. Each indicator reflects a distinct aspect of visualization effectiveness, including visual efficiency, cognitive load, and task performance, making equal weighting a simple and transparent integration strategy.

A preliminary sensitivity analysis was conducted by varying the weights within reasonable ranges, and the overall trends remained consistent, indicating that the VEI is relatively robust to moderate weight changes. Future work will explore data-driven weighting methods, such as analytic hierarchy process (AHP) or regression-based approaches, to further improve the model.

The VEI integrates visual attention efficiency, cognitive load, and task performance into a single quantitative measure, enabling direct comparison between visualization designs and supporting data-driven optimization.

4 Experiments and Comparative Analysis

4.1 Experimental Setup

The experimental data were collected through a controlled eye-tracking study conducted in a laboratory environment. A total of 60 participants (30 experts and 30 novices) completed task-oriented evaluations based on three types of information visualizations: tree diagrams, heatmaps, and Sankey diagrams. All participants had normal or corrected-to-normal vision.

Three design conditions were evaluated: Baseline, Traditional, and Proposed. The Baseline condition represents an unoptimized design without intentional enhancement of layout, color, or information hierarchy. The Traditional condition follows common visualization design principles, including basic color encoding and standard visual hierarchy. The Proposed condition is optimized based on the proposed evaluation framework to improve attention guidance and information comprehension. All conditions were constructed using the same datasets and tasks, differing only in design strategy.

Each visualization type included multiple stimuli with comparable complexity. Tasks covered information identification, comparison, and interpretation, with controlled difficulty to ensure balanced cognitive load. The presentation order was randomized across participants, and a short practice session was conducted before the experiment to reduce learning effects.

Eye movement data were recorded using Tobii Pro Glasses 3 (100 Hz). Calibration was performed before each session, and the experiment was conducted under controlled lighting conditions. Fixation detection was implemented using the I-VT algorithm.

The experimental procedure included calibration, practice, and formal testing. Evaluation was based on eye-tracking metrics (FFD, TFD, Scanpath Length, Entropy) and task performance (Accuracy Rate). Statistical significance was tested using one-way ANOVA ($p < 0.05$).

4.2 Main Experimental Results

The main experimental results are summarized in Table 1, which presents a comparative analysis of different visualization design methods based on the selected evaluation metrics. Three design conditions were evaluated: Baseline (unoptimized design), Traditional (designed following common visual principles), and Proposed (optimized using the proposed evaluation model).

Table 1. Performance comparison of different methods

Method	FFD (ms) ↓	TFD (ms) ↓	Scanpath Length ↓	Entropy ↑	Accuracy (%) ↑
Baseline	517 ± 48	4125 ± 320	1798 ± 210	1.24 ± 0.08	73.2 ± 6.5
Traditional	438 ± 41	3470 ± 285	1495 ± 175	1.36 ± 0.07	80.6 ± 5.8
Proposed	**326 ± 36**	**3015 ± 260**	**1218 ± 160**	**1.58 ± 0.06**	**88.9 ± 4.9**

As shown in Table 1, the proposed method consistently outperforms the baseline and traditional approaches across all metrics. Specifically, the Proposed design achieves a substantially lower FFD (326 ms) compared with the Baseline (517 ms) and Traditional (438 ms), indicating more efficient visual guidance. Similarly, the TFD and scanpath length are significantly reduced, suggesting lower cognitive load and more coherent visual exploration behavior.

In contrast, the Proposed method yields higher entropy (1.58) and accuracy (88.9%), reflecting a more balanced distribution of visual attention and improved task comprehension. The performance improvements across all metrics are statistically significant ($p < 0.05$), confirming the effectiveness and robustness of the proposed evaluation model.

4.3 Discussion

The results show that the proposed method improves visualization effectiveness compared with baseline and traditional designs. Reduced FFD and scanpath length indicate faster information localization and more efficient visual guidance, while decreased TFD suggests lower cognitive load. The increase in entropy reflects a more distributed allocation of visual attention across regions, which may support more comprehensive information exploration depending on task requirements.

The higher accuracy rate further indicates improved user comprehension. These findings support the use of eye-tracking metrics for visualization evaluation. The proposed VEI model provides a quantitative framework linking visual behavior with task performance, offering practical value for design optimization.

5 Conclusion

This study proposes an eye-tracking-based model for evaluating information visualization effectiveness by integrating eye movement metrics with task performance into a composite index (VEI). Experimental results show that the model effectively distinguishes design approaches and provides insights into visual efficiency, cognitive load, and user comprehension. Compared with traditional subjective methods, it improves reliability and reproducibility.

However, this study has limitations. The participants were limited to students, and the experiments focused on static visualizations. Future work will extend the model to more diverse user groups and interactive scenarios, and explore multimodal data such as EEG to further enhance evaluation accuracy.

References

1. Q. Chen, T. Doan, et al., "Event-Based Eye Tracking: Methods and Applications," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2025. arXiv: 2504.09960
2. A. Rehman, I. Heldal, and J. C. W. Lin, "VisiTrail: A Cognitive Visualization Tool for Time-Series Analysis of Eye Tracking Data from Attention Game," arXiv preprint arXiv:2508.20522, 2025.
3. E. Mehmedova, S. Berrezueta-Guzman, and S. Wagner, "iTrace: Click-Based Gaze Visualization on the Apple Vision Pro," arXiv preprint arXiv:2508.12268, 2025.
4. R. Khan, et al., "Eye-Tracking Technologies: Theory, Methods and Applications," Applied Sciences, vol. 14, no. 5, pp. 1789, 2024. doi: 10.3390/app14051789
5. M. J. Dunn, R. G. Alexander, O. M. Amiebenomo, et al., "Minimal Reporting Guideline for Research Involving Eye Tracking (2023 Edition)," Behavior Research Methods, vol. 56, no. 5, pp. 4351–4357, 2024. doi: 10.3758/s13428-023-02187-1
6. W. Fuhl, G. Kasneci, and E. Kasneci, "TEyeD: Over 20 Million Real-World Eye Images with Pupil, Eyelid, and Iris 2D and 3D Segmentations, 2D and 3D Landmarks, 3D Eyeball, Gaze Vector, and Eye Movement Types," arXiv preprint arXiv:2107.13937, 2021.
7. K. Kurzhals, M. Burch, and D. Weiskopf, "Eye Tracking in Visualization: Beyond Gaze Distributions and Scanpaths," IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 5, pp. 2448–2461, 2021. doi: 10.1109/TVCG.2021.3067779