Journal of Educational Technology and Innovation
https://jeti.thewsu.org/index.php/cieti

Article

Visualizing Expert-Novice Reading Strategies in Classical Chinese Poetry: An Application of Hidden Markov Models

Feng Xiong [1,†], Anran Ma [1,2,†], Haixia Yuan [3,*], Ming Li [4,*] and Yanli Chen [5]

1 National Engineering Research Center of Educational Big Data, Central China Normal University, Wuhan 430079, China
2 School of Psychology, Central China Normal University, Wuhan 430079, China
3 School of International Cultural Exchange, Central China Normal University, Wuhan 430079, China
4 School of Journalism and Communication, Hubei University of Economics, Wuhan 430205, China
5 School of Wisdom and Health, Chongqing Vocational University of Electronic Science and Technology, Chongqing 400716, China
* Correspondence: yuanhaixia@ccnu.edu.cn (H.Y.); liming2@hbue.edu.cn (M.L.)
† These authors contributed equally to this work.

How To Cite: Xiong, F., Ma, A., Yuan, H., Li, M., & Chen, Y. (2026). Visualizing Expert-Novice Reading Strategies in Classical Chinese Poetry: An Application of Hidden Markov Models. *Journal of Educational Technology and Innovation*, *8*(1), 25–39. https://doi.org/10.61414/2msjhn78

Received: 30 December 2025
Revised: 10 February 2026
Accepted: 12 March 2026
Published: 31 March 2026

Abstract: Classical Chinese Poetry represents a unique intersection of linguistic art and cultural heritage, yet the spatiotemporal dynamics of its reading process remain difficult to observe using traditional methods. While previous studies have relied on static eye-tracking metrics, this study applies a data-driven machine learning approach to visualize the distinct reading strategies of learners. Using eye-movement analysis with hidden Markov models (EMHMM), we recorded and modeled the gaze patterns of 35 participants (experts vs. novices) as they read a classical poem. The results successfully visualized two distinct spatiotemporal patterns through representative case modeling: the novice exemplar exhibited a "text-bound, local scanning" strategy characterized by high self-loop probabilities, reflecting high cognitive load; in contrast, the expert exemplar employed a "schema-driven, global scanning" strategy, efficiently integrating imagery. Furthermore, the application of EMHMM revealed significant differences in the late-stage processing of "Poetry Eye" (expressive words), highlighting the gap in aesthetic integration. Critically, this study demonstrates the utility of EMHMM as a visualization tool for educational assessment. We propose that these quantified spatiotemporal patterns demonstrate the potential to serve as prototypical computational biomarkers, providing empirical evidence for the design of AI-driven intelligent tutoring systems (ITS) that offer gaze-contingent scaffolding.

Keywords: eye-tracking; hidden Markov models (HMM); classical Chinese poetry; expert-novice paradigm; learning analytics; intelligent tutoring systems

1. Introduction

Classical Chinese Poetry, regarded as the crystallization of linguistic art and cultural heritage, demands a highly sophisticated cognitive process for appreciation. Unlike expository texts, poetry comprehension requires readers to transcend surface-level decoding and engage in aesthetic simulation and emotional resonance (Jacobs, 2018; Tang, 2020). According to Kintsch's (2018) Construction-Integration Model, successful appreciation hinges on the reader's ability to integrate fragmented imagery into a coherent mental model. However, recent studies highlight that this process is heavily dependent on mental imagery generation and cognitive fluency, which imposes a significant load on novice learners (Lei & He, 2025). When readers lack the necessary prior knowledge schemas to support this high-element interactivity, their working memory is easily overwhelmed, leading to a breakdown in aesthetic integration (Hattan et al., 2024; Sweller, 2024).

Despite the importance of poetry education, observing how learners navigate these challenges remains difficult. Traditional research has largely relied on subjective interpretation. While eye-tracking technology has been introduced, most studies utilize static metrics, such as total fixation counts or heatmaps (Drusch et al., 2014). As noted by Raschke et al. (2013), these aggregated measures fail to visualize the spatiotemporal dynamics of reading—they reveal where a student looked, but obscure the sequence of attention shifts. This limitation is critical, as recent neurocognitive evidence suggests that the core difference between experts and novices lies in the dynamic interplay between top-down and bottom-up processing over time (Flower & Pylkkänen, 2024; Bandiera et al., 2025). Consequently, the distinct scanning strategies that differentiate experts from novices have not been fully visualized or quantified using static methods.

To address this methodological limitation, this study applies eye-movement analysis with Hidden Markov Models (EMHMM) to visualize the cognitive strategies of poetry reading. EMHMM is a data-driven technique capable of uncovering latent cognitive states and transition probabilities (Wang & Zhan, 2025; Hsiao et al., 2021). An earlier, preliminary version of this research was presented at the 2023 International Conference on Intelligent Education and Intelligent Research (IEIR) (Xiong et al., 2023). While the conference paper established the basic spatial eye-movement differences using clustering, the current manuscript significantly extends this work by implementing advanced temporal and network analyses to visualize the underlying cognitive state-switching dynamics, and by re-contextualizing these findings within the Expert-Novice paradigm. By applying this tool to contrast the scan paths of high-prior-knowledge (expert) and low-prior-knowledge (novice) readers, we aim to provide a dynamic visualization of the “Global vs. Local” processing strategies.

Specifically, to address the methodological gaps in dynamic observation, the main aim of this study is to apply EMHMM as a visualization tool to quantify the cognitive differences between expert and novice readers. This study focuses on three key objectives:

- Visualization of Strategies: To apply EMHMM to visualize the distinct hidden states (ROIs) and spatiotemporal transition dynamics characterizing expert versus novice reading pathways.
- Processing of “Poetry Eye”: To investigate how prior knowledge schemas influence the temporal processing of the “Poetry Eye” (expressive words) through dynamic modeling.
- Educational Application: To explore the potential of these visualized patterns as potential computational biomarkers for AI-driven process assessment in intelligent tutoring systems.

2. Related Work

2.1. Cognitive Mechanisms in Poetry Reading

Reading Classical Chinese Poetry is a complex cognitive activity that transcends simple text decoding. According to the framework of Neurocognitive Poetics, poetry comprehension involves not just linguistic processing but also aesthetic appreciation and emotional resonance (Jacobs, 2018; Aryani, 2018; Stopel, 2018). Recent studies have further highlighted that the aesthetic experience of Tang poems is deeply rooted in cognitive fluency and the successful retrieval of artistic conceptions (Tang, 2020).

To understand this multifaceted process, Kintsch’s Construction-Integration (CI) Model provides a robust theoretical framework (Kintsch, 2018). The model suggests that comprehension occurs in two phases: the construction of a textbase (deriving literal meaning from the text) and the integration of a situation model (forming a coherent mental representation) (Aryadoust, 2019; Zhang et al., 2023). In the context of poetry, readers must move beyond surface-level construction to integrate abstract imagery and narrative voices into a unified mental model (Chen et al., 2023). This process is further supported by Dual-Coding Theory, which posits that poetry invokes both verbal systems (language processing) and non-verbal systems (mental imagery generation) (Sadoski, 2018). Empirical evidence from eye-tracking studies confirms that literary reading triggers active mental simulations, where eye movements reflect the reader’s internal visualization of the text’s imagery (Magyari et al., 2020; Mak & Willems, 2019). Specifically, for Classical Chinese Poetry, the generation of mental imagery has been found to be a critical predictor of both comprehension and aesthetic appreciation (Lei & He, 2025).

However, this dual processing imposes significant demands on working memory. Cognitive Load Theory suggests that working memory is a limited resource that can be easily depleted by high intrinsic load (Puma & Tricot, 2019; Sepp et al., 2019). When learners lack the necessary prior knowledge or schemas to process high-element interactivity—such as the obscure vocabulary or complex syntax often found in classical poetry—their working memory becomes overwhelmed (Chen et al., 2018; Greenberg & Zheng, 2022). Consequently, novice readers often struggle to reach the stage of aesthetic integration, getting stuck in the decoding phase, as their cognitive resources are exhausted by the effort of basic text processing.

2.2. Eye-Tracking Technology and The Expert-Novice Paradigm

Eye-tracking methodology, based on the "Eye-Mind Assumption," serves as an objective window into these underlying cognitive processes, positing that where we look directly reflects what we are processing (Just & Carpenter, 1980). Research employing the Expert-Novice Paradigm has consistently demonstrated that domain expertise fundamentally alters visual scanning strategies across diverse tasks.

Schema Theory provides the primary explanation for these differences. According to Anderson (2018), experts possess sophisticated mental structures (schemas) that allow them to process information in a "top-down" manner. Recent reviews emphasize that the activation of prior knowledge significantly reduces cognitive load by enabling readers to predict and integrate information more efficiently (Hattan et al., 2024; Smith et al., 2021). This top-down mechanism enables holistic processing, where experts scan globally to extract meaning from large perceptual spans. Empirical studies confirm that top-down factors, such as semantic expectation, dominate eye movements in proficient readers (Flower & Pylkkänen, 2024; Bandiera et al., 2025). In contrast, novices, lacking these robust schemas, rely on "bottom-up" analytic processing, attending to local details sequentially to decipher meaning (Breadmore & Carroll, 2018).

Furthermore, the Information Reduction Hypothesis proposes that expertise is characterized by the ability to selectively ignore redundant information and focus solely on task-relevant areas (Haider & Frensch, 1999). This hypothesis has been robustly supported by recent cross-disciplinary research. For instance, in medical image diagnosis, experts (e.g., radiologists) demonstrate efficient, global scan paths, focusing rapidly on abnormalities while ignoring healthy tissue (Ivy et al., 2023; Lee et al., 2023). Similarly, in visual art appreciation, experts exhibit broader scanning patterns focused on composition and style, whereas novices fixate heavily on salient local features (Yi et al., 2021; Pelowski et al., 2018). These findings suggest a universal principle: expertise facilitates a shift from local, feature-based scanning to global, meaning-based appreciation. This study aims to verify whether this phenomenon applies specifically to the appreciation of Classical Chinese Poetry.

2.3. Machine Learning in Learning Analytics

While traditional eye-movement measures (e.g., fixation count, heatmaps) provide aggregated statistics, they often fail to capture the spatiotemporal dynamics of the reading process. As noted by Drusch et al. (2014) and Raschke et al. (2013), static metrics cannot reveal how attention shifts over time or in what sequence information is processed. Although scan path comparison algorithms like MultiMatch (Dewhurst et al., 2012) and String-Edit Distance (Eraslan et al., 2015) attempt to address this, they are often computationally expensive and struggle to generalize across individuals (French et al., 2017).

To address these limitations, researchers in Learning Analytics have increasingly turned to machine learning approaches, specifically Hidden Markov Models (HMM) (Kui et al., 2022; Raković et al., 2023). In the domain of visual cognition, eye-movement analysis with hidden Markov models (EMHMM) has emerged as a powerful tool to quantify individual scan paths (Hsiao et al., 2021; Chuk et al., 2020). Unlike standard methods that require pre-defined Areas of Interest (AOIs), EMHMM is data-driven, capable of uncovering latent cognitive states (ROIs) and the transition probabilities between them (Hsiao & Chan, 2019; Chuk et al., 2017). Recent applications have successfully used EMHMM to reveal cognitive strategy switching (Wang & Zhan, 2025), predict reading performance (Zheng et al., 2022), and detect mind-wandering (Lee et al., 2021).

This shift towards Process Mining allows for a more granular assessment of learning strategies (Grub et al., 2020; Jankowska et al., 2018). Recent studies in AI-driven assessment suggest that such computational modeling can serve as the foundation for intelligent tutoring systems (ITS), offering real-time diagnosis of student cognitive states (D'Mello & Graesser, 2023; Azevedo et al., 2022). By integrating multimodal data, including eye gaze, these systems can provide adaptive scaffolding to enhance learning outcomes (Wang et al., 2021; Chettaoui et al., 2023).

In summary, while existing literature has established the cognitive complexity of poetry reading and the utility of eye-tracking in expertise research, significant gaps remain. Most current studies continue to rely on static, aggregated metrics (e.g., fixation counts), which obscure the temporal dynamics of how readers construct meaning over time. Furthermore, although machine learning approaches like HMM have shown promise in general visual cognition, their application to the aesthetic appreciation of Chinese classical poetry—a domain requiring the unique integration of imagery and emotion—remains underexplored. Specifically, there is a lack of research that not only quantifies these dynamic patterns but also translates them into actionable visual insights for educational intervention. Consequently, the present study addresses this necessity by applying EMHMM to visualize these invisible cognitive strategies, thereby bridging the gap between cognitive theory and AI-driven pedagogical practice.

3. Methodology

3.1. Participants and Experimental Design

A total of 45 undergraduate and graduate students were recruited via campus advertisements from Central China Normal University. To ensure a distinct differentiation in expertise, an extreme group design was employed. Based on the aggregate scores from the Experience on Chinese Poem Questionnaire, participants with intermediate scores were excluded. The final valid sample (N = 35) was stratified into two distinct cohorts: the high prior knowledge group (Experts, n = 13), comprising participants with scores > 10; and the low prior knowledge group (Novices, n = 22), comprising participants with scores < 9. To guarantee the validity of oculomotor data, strict inclusion criteria were applied: participants were required to have normal or corrected-to-normal vision (visual acuity > 1.0), be right-handed, and have no history of neurological, psychiatric, or reading disorders (e.g., dyslexia).

3.2. Instruments and Stimuli

The study utilized three psychometric instruments and one visual stimulus. (1) The Experience on Chinese Poem Questionnaire assessed the breadth of participants' exposure to classical literature (Cronbach's α = 0.79). (2) The Learning Outcome Questionnaire evaluated comprehension through four open-ended dimensions: literal translation, aesthetic appreciation of "Poetry Eye", stylistic analysis, and emotional interpretation. Scoring was conducted by two independent raters with high inter-rater reliability (Cohen's κ = 0.85). (3) A single-item Perceived Difficulty Scale (1–9) measured subjective cognitive load. (4) Stimuli: The experimental stimulus was a seven-character quatrain (Tang Dynasty poem, see Figure 1) presented in a traditional vertical layout with black Calligraphy font (KaiTi) on a white background. It is a seven-character quatrain from the Tang Dynasty displayed in traditional calligraphy font. The critical "Poetry Eye" (character "Laugh"/"笑") is located in the third line. The characters were rendered at a font size of 32 pixels, subtending a visual angle of approximately 1.2° per character at a 75 cm viewing distance, ensuring that foveal vision encompassed individual characters while minimizing lateral interference.

春

一春略无十日晴，处处浮云将雨行。

野田春水碧于镜，人影渡傍鸥不惊。

桃花嫣然出篱笑，似开未开最有情。

茅茨烟暝客衣湿，破梦午鸡啼一声。

Figure 1. The experimental stimulus of Classical Chinese Poetry presented to participants. The poem, titled "Spring", consists of four couplets. Sentence-by- sentence translation: (1) Barely ten sunny days throughout spring; (2) Clouds drift everywhere, carrying the rain; (3) Spring waters in wild fields are greener than a mirror; (4) Gulls by the ferry are not startled by passing human shadows; (5) Peach blossoms smile (xiao) brightly through the hedge; (6) They are most evocative when half-opened; (7) Mists shroud the thatched hut as the traveler's clothes dampen; (8) A single noon-day cock-crow breaks the sleeping dream. The character "笑" (highlighted in sentence 5) serves as the "Poetry Eye" (shiyan), a key expressive focus in this study.

3.3. Apparatus and Data Acquisition

Eye movements were recorded using an EyeLink 1000 desktop eye tracker (SR Research Ltd., Ottawa, ON, Canada) configured in Tower Mount mode. The system sampled data at a frequency of 500 Hz, achieving a high spatial resolution (RMS) of less than 0.01° and an average tracking accuracy within the range of 0.25° to 0.5°. The stimuli were displayed on a 19-inch CRT monitor (refresh rate: 100 Hz; resolution: 1280 × 1024 pixels), providing a linear physical canvas for gaze mapping. A chin-and-forehead rest was employed to stabilize the participant's head, minimizing motion artifacts. The experimental procedure was programmed and executed using SR Research Experiment Builder software.

3.4. Procedure

The experiment was conducted in a sound-attenuated, dimly lit laboratory. The session commenced with the informed consent process and the prior knowledge assessment. Subsequently, participants underwent a standard

9-point calibration and validation procedure. The validation acceptance threshold was set strictly at an average gaze error of <0.5° and a maximum error of <1.0°. A drift check was performed immediately prior to the trial to correct for any postural shifts. The poem was then presented for a fixed duration of 60 s, a timeframe determined by pilot testing to be sufficient for deep appreciation. Participants were instructed to read silently for comprehension and aesthetic appreciation. Following the eye-tracking task, they completed the post-test questionnaires. The total session duration was approximately 15 min.

3.5. Data Analysis

3.5.1. Participants and Data Pre-processing

Data analysis was performed in three distinct stages to systematically reveal the cognitive processing differences between experts and novices. A total of 35 valid participants were included in the final analysis. Based on their scores from the prior knowledge assessment, participants were divided into a high prior knowledge group (experts, $n = 13$) and a low prior knowledge group (novices, $n = 22$). Prior to statistical analysis, raw eye-movement data underwent strict pre-processing to ensure data quality. Specifically, fixations shorter than 80 ms or longer than 800 ms were excluded as outliers, ensuring that the subsequent metrics reflected genuine cognitive processing rather than artifacts.

3.5.2. Statistical Analysis

The quantitative analysis comprised the first two stages of the study. In the first stage, independent samples t-tests were conducted to compare learning outcomes and perceived difficulty between the two groups ($N = 35$), providing a behavioral baseline for performance differences. In the second stage, we focused on static eye-movement metrics at two granularities: the global level (the entire poem) and the local level (specifically the "Poetry Eye", i.e., the character "Laugh (笑)"). We selected First Fixation Duration to index early lexical processing difficulty and Total Fixation Count to reflect the depth of semantic integration. Independent samples t-tests were performed for these metrics to assess group differences. Additionally, Cohen's d was calculated to estimate the effect size for all significant findings, with the significance level set at $\alpha = 0.05$.

3.5.3. Cognitive Strategy Modeling (EMHMM)

In the third stage, we employed the Eye Movement analysis with Hidden Markov Models (EMHMM) toolbox to qualitatively visualize dynamic spatiotemporal strategies. To extract prototypical scan paths without overfitting the model to the unequal sample sizes ($n = 13$ vs. $n = 22$), we adopted a representative sampling approach rather than modeling the entire group averages. To extract prototypical scan paths without overfitting the model to the unequal sample sizes (*n* = 13 vs. *n* =2 2), we adopted a purposive sampling approach for a qualitative case study. Specifically, the individual visualized eye-movement trajectories of all participants were comprehensively reviewed via visual inspection. Subject 5 (Expert) and Subject 23 (Novice) were selected because their scan paths exhibited the most distinct and typical spatiotemporal features of their respective cohorts, conceptually serving as the "cluster centers" or prototypical exemplars. This subset analysis allowed us to qualitatively visualize the distinct "global schema-driven" versus "local text-bound" scan paths. This subset analysis allowed us to visualize the distinct "global schema-driven" versus "local text-bound" scan paths characterizing each expertise level by illustrating the specific transition probabilities between Regions of Interest (ROIs).

Model Formulation of EMHMM. We modeled the reading process as a temporal sequence of eye fixations generated by hidden cognitive states. Specifically, the EMHMM assumes that the reader's attention shifts between a set of Regions of Interest (ROIs), where each ROI is associated with a Gaussian emission reflecting the spatial distribution of fixations. The transitions between these ROIs represent the underlying cognitive strategy. We modeled the eye-movement trajectory of each participant as a temporal sequence of fixations. Let $O = \{o_1, o_2, \ldots, o_T\}$ denote the sequence of eye fixations for a single trial, where T is the total number of fixations. Each observation $o_t = (x_t, y_t)$ represents the 2D coordinates of the fixation at time t on the screen.

The HMM assumes that these observed fixations are generated by a sequence of hidden cognitive states $H = \{h_1, h_2, \ldots, h_T\}$, where each state $h_t \in \{1, \ldots, K\}$ and corresponds to a specific ROI focused on by the reader. The model is characterized by three key components:

(1) Prior Probabilities: The probability distribution of the initial fixation landing on a specific ROI:

$$\pi_i = P(h_1 = i), \quad 1 \le i \le K \tag{1}$$

(2) Transition Probabilities: The probability of the eyes moving from ROI i at time $t-1$ to ROI j at time t, reflecting the sequential reading strategy:

$$A_{ij} = P(h_t = j \mid h_{t-1} = i), \quad 1 \leq i, j \leq K \tag{2}$$

This transition matrix captures the dynamic scanning patterns, such as sequential reading (linear transitions) or regressions (backward transitions).

(3) Emission Probabilities: The probability of observing a fixation coordinate o_t given the current hidden state h_t. In the EMHMM, each ROI is modeled as a 2D Gaussian distribution rather than a static box. Thus, the emission probability is defined as:

$$\mathrm{P}(o_t \mid h_t = i) = \mathcal{N}(o_t; \mu_i, \Sigma_i) \tag{3}$$

where $\mathcal{N}$ denotes the probability density function of a multivariate Gaussian distribution, μ_i represents the center of the i-th ROI and Σ_i represents the spatial dispersion (covariance) of fixations within that ROI.

Parameter Estimation and Model Selection. The model parameters, including the transition probability matrix and emission densities, were estimated using the Variational Bayesian Expectation-Maximization (VB-EM) algorithm. To determine the optimal number of hidden states (ROIs) for each group, we utilized the clustering methodology based on log-likelihood variation, ensuring that the model complexity balanced interpretability with data fit. The parameters of the HMM, denoted as $\lambda = \{\pi, A, \mu, \Sigma\}$, were estimated using the Variational Bayesian Expectation-Maximization (VB-EM) algorithm. This method is preferred over the standard Maximum Likelihood Estimation (MLE) because it avoids overfitting and effectively handles the complexity of individual differences in eye-movement data. For each participant, we trained individual HMMs to capture their unique reading patterns. To determine the optimal number of hidden states (ROIs) K, we utilized the Bayesian Information Criterion (BIC) to balance model complexity and goodness of fit.

Visualization of Spatiotemporal Strategies. Finally, to reveal the distinct reading patterns, we visualized the estimated models as cognitive network graphs. By mapping the transition probabilities onto the poem's layout, we reconstructed the prototypical "global schema-driven" scan paths for experts and "local text-bound" paths for novices, thereby making the invisible cognitive dynamics observable.

4. Results

4.1. Verification of Grouping Based on Prior Knowledge

To ensure the validity of the comparative design, participants were stratified into two distinct cohorts based on their scores from the Experience on Chinese Poem Questionnaire (total score = 21). Data screening resulted in a valid sample distribution: the high prior knowledge group (experts) comprised 13 participants with scores exceeding 10, while the low prior knowledge group (novices) consisted of 22 participants with scores below 9. This polarized distribution confirms a clear differentiation in expertise levels between the two groups, providing a solid foundation for the subsequent expert-novice comparison.

4.2. Behavioral Results of Learning Outcomes and Cognitive Load

Behavioral data analyses revealed significant differences in both learning performance and subjective difficulty. Regarding learning outcomes, the expert group (M = 18.77, SD = 1.03) demonstrated significantly superior performance compared to the novice group (M = 16.55, SD = 1.53). An independent samples t-test confirmed this disparity (t (33) = −2.27, p = 0.03, η^2= 0.14), indicating that prior knowledge significantly facilitated the comprehension and appreciation of the poem.

Conversely, regarding subjective cognitive load (perceived difficulty), the novice group reported significantly higher difficulty levels (M = 5.82, SD = 1.53) compared to the expert group (M = 4.69, SD = 1.03). The statistical analysis yielded a significant difference (t (33) = -2.35, p = 0.025, η^2= 0.14). These behavioral findings suggest that while experts achieved better comprehension with lower perceived effort, novices struggled with higher cognitive load and lower performance.

4.3. Eye-Movement Metrics Analysis

To investigate the time course of cognitive processing and its relationship with learning outcomes, eye-movement analyses were conducted on three levels: a holistic analysis of the entire poem, a focal analysis of the "Poetry Eye" (the specific character "Laugh"/"笑"), and a correlation analysis.

4.3.1. Global Processing Patterns

At the global level (the entire poem), the analysis revealed a similar distribution of cognitive resources between groups (see Figure 2). Although the expert group showed a slightly higher mean First Fixation Duration and Total Dwell Time compared to the novice group, these differences did not reach statistical significance ($p > 0.05$). Similarly, no significant difference was found in the overall Fixation Count ($p > 0.05$). These findings indicate that in terms of basic reading load and overall attention allocation, there is no fundamental disparity between experts and novices when reading the poem as a whole.

4.3.2. Focal Analysis of the "Poetry Eye"

As illustrated in Figure 2, we conducted a comparative analysis of key eye-movement metrics between the novice and expert groups at two distinct levels: the macroscopic level (whole poem) and the microscopic level (the specific ROI of the character "笑"/"Laugh").

At the whole-poem level (Figure 2, top row), we examined First Fixation Duration, Fixation Count, and Total Dwell Time. The statistical analysis revealed no significant differences between the two groups across these metrics ($p > 0.05$). This lack of divergence suggests that, in terms of overall cognitive effort and temporal allocation throughout the entire poem, novices and experts performed similarly.

In contrast, the analysis of the emotional keyword "笑" (Figure 2, bottom row) yielded distinct patterns. Specifically, the novice group demonstrated a significantly higher Fixation Count compared to the expert group ($p < 0.05$). This disparity indicates that novices employed a "selective deep processing" strategy, allocating disproportionately greater visual attention to emotionally salient text compared to the more distributed attention pattern observed in experts.

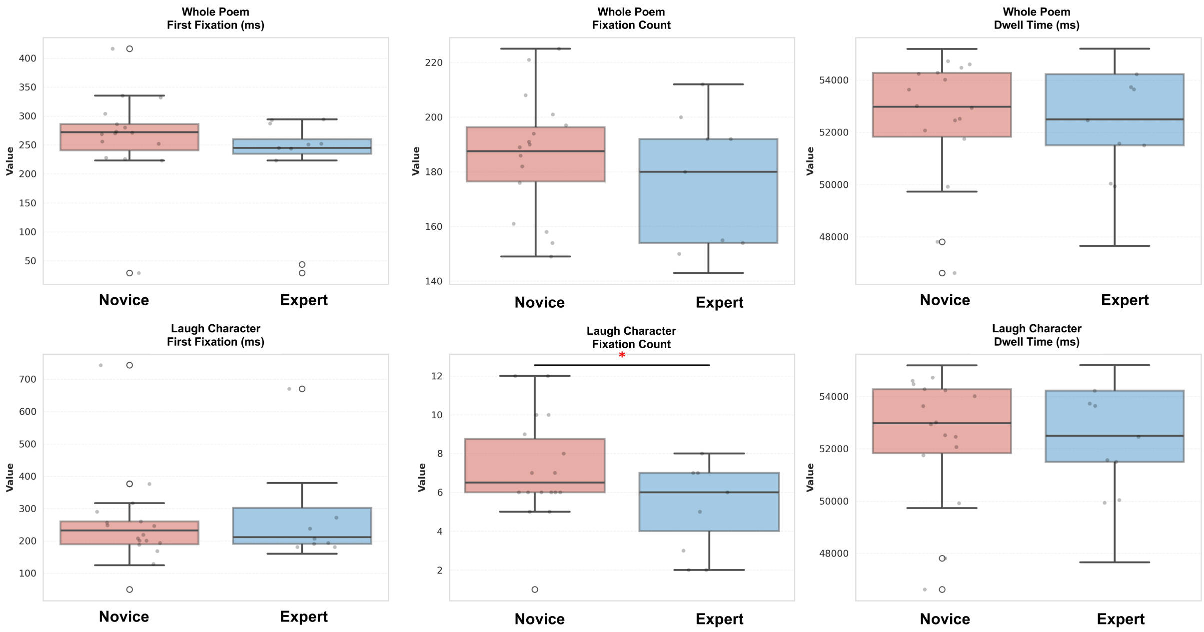

Figure 2. Statistical comparison of key eye-movement metrics between novice and expert groups. The top row displays metrics for the whole poem (First Fixation Duration, Fixation Count, and Total Dwell Time), while the bottom row compares metrics for the specific ROI of the character "笑" (Laugh). Note: Box plots display the median (center line), interquartile range (box), and whiskers (1.5 × IQR). Individual data points are overlaid with jitter. Asterisks indicate statistical significance: * $p < 0.05$, ns = not significant.

Further analysis focused on the critical ROI, the "Poetry Eye" (character "Laugh"). Cognitive processing indicators were categorized into early processing measures (First Fixation Duration) and late processing measures (Total Fixation Count).

(1) Early Processing: No significant difference was found in First Fixation Duration between experts (M = 249.09, SD = 134.58) and novices (M = 268.31, SD = 136.11; t (33) = −0.41, p = 0.68). This indicates that the initial lexical recognition of the character—a common word in Chinese—was automated and equally efficient for both groups.

(2) Late Processing: Significant divergences emerged in late-stage indicators. For Total Fixation Count, novices (M = 7.14, SD = 2.58) exhibited significantly more fixations than experts (M = 5.38, SD = 2.10; t (33) = 2.06, p = 0.047). These results imply that while lexical access was similar, novices experienced significantly greater difficulty in integrating this character into the poetic context.

4.3.3. Correlation between Oculomotor Metrics and Learning Outcomes

To explore whether specific eye-movement patterns on the “Poetry Eye” directly predicted learning performance, Pearson correlation analyses were conducted between eye-movement metrics (First Fixation Duration, Total Fixation Count) and scores on the “word appreciation” sub-questionnaire. The results indicated no statistically significant correlations (p > 0.05). This suggests that while group-level differences exist, the relationship between specific local gaze behaviors and explicit test scores may be non-linear or mediated by broader scanning strategies, rather than determined solely by the processing of a single character.

4.4. Spatial Features: A Qualitative Case Study of Prototypical Scan paths

To elucidate the spatiotemporal dynamics of reading, we employed Eye Movement Hidden Markov Models (EMHMM) on a representative subset of participants. As visualized in Figure 3, the analysis revealed two distinct reading strategies.

the expert group (Figure 3, left panel) exhibited a “global integration pattern.” This strategy is characterized by multiple bidirectional transitions (e.g., the interaction between ROI 1 and ROI 2) and cyclic loops (ROI 3 → ROI 4 → ROI 3). These non-linear movements indicate an active search for meaning integration, suggesting that experts continuously synthesize information across distant semantic regions of the poem.

In contrast, the novice group (Figure 3, right panel) demonstrated a “linear scanning pattern.” Their scan paths were dominated by simple, sequential transitions (e.g., a unidirectional shift from ROI 2 to ROI 1) and underlying self-loops on individual characters. This pattern reflects a text-bound, surface-level decoding process, where high cognitive load limits the reader to local processing with minimal global integration.

It is crucial to note that the following EMHMM spatial networks and transition matrices are presented as a qualitative case-study visualization. While these prototypical exemplars were carefully selected to represent the broader behavioral tendencies of their respective groups, these network graphs aim to dynamically illustrate the underlying cognitive state-switching strategies rather than provide absolute quantitative population parameters.

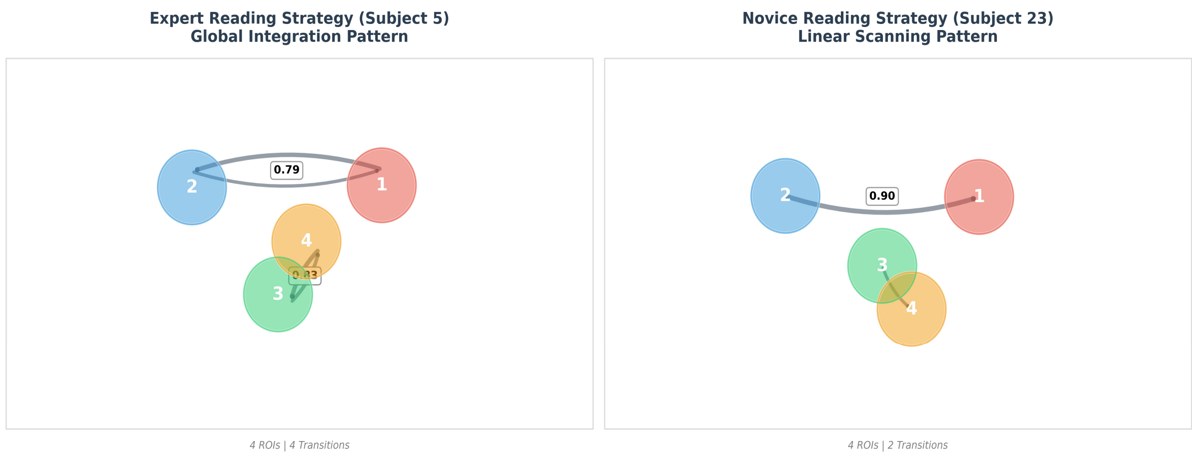

Figure 3. Comparison of eye-movement strategies between expert and novice readers using Eye Movement Hidden Markov Models (EMHMM). The left panel illustrates the representative expert strategy (subject 5), and the right panel shows the novice strategy (subject 23). Note: Regions of Interest (ROIs) are represented by semi-transparent colored circles. Curved arrows indicate transition probabilities between ROIs, with line thickness proportional to the probability value. For visual clarity, only key transitions with probabilities > 0.40 are displayed, and self-loops are omitted.

4.4.1. The "Text-Bound" Strategy in the Novice Case

As illustrated in the clustering results (Figure 4), the novice exhibited a distinctly "constrained" spatial pattern that aligns with a bottom-up processing strategy.

Spatially (Figure 4, left panel), the identified Regions of Interest (ROIs) were characterized by smaller covariance ellipsoids that adhered strictly to the linear syntactic layout of the verses. Notably, the analysis revealed a high-density ROI specifically centered around the "Poetry Eye" (the character "笑"/"Laugh") in the third line. This intense focal attention suggests that novices prioritize decoding emotionally salient lexical units over integrating broader semantic fields.

Temporally (Figure 4, right panel), the transition network further confirms this "text-bound" trajectory. Unlike the cyclic loops observed in experts, the novice scan path is dominated by localized connections and unidirectional paths (e.g., the linear shift from ROI 2 to ROI 1). This lack of global integration indicates that novices allocate cognitive resources primarily to sequential decoding, reflecting the higher cognitive load associated with lower domain knowledge.

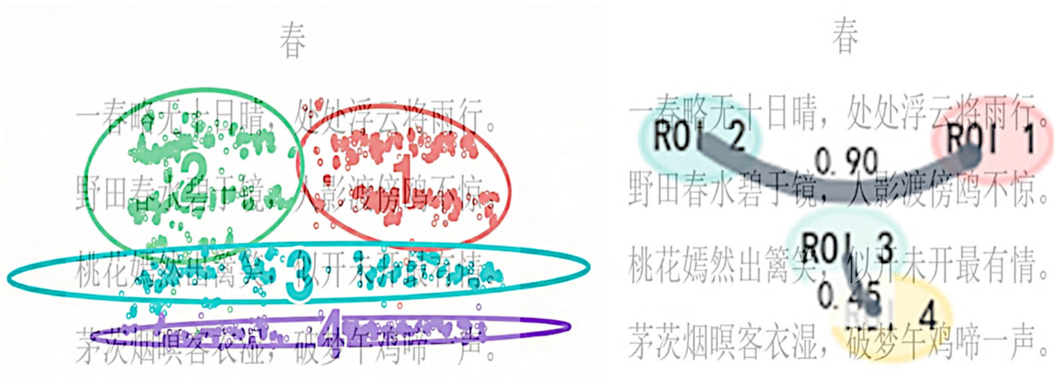

Figure 4. Visualization of the reading strategy of a representative novice (subject 23) modeled by EMHMM. The **left panel** illustrates spatial distribution of fixations (dots) and resulting regions of interest (ROIs). Distinct colors denote different ROIs: Red (ROI 1), Green (ROI 2), Cyan (ROI 3), and Purple (ROI 4). The **right panel** illustrates transition network showing probabilities between identified ROIs. Note: The layout illustrates a linear progression, with ROIs aligned to text structure (e.g., ROI 2 → ROI 1). The specific focus on the character "Laugh" (笑) is captured within ROI 3.

4.4.2. The "Holistic" Strategy in the Expert Case

In distinct contrast to the novice pattern, the expert displayed a "holistic" spatial strategy that transcends the linear constraints of the text (see Figure 5).

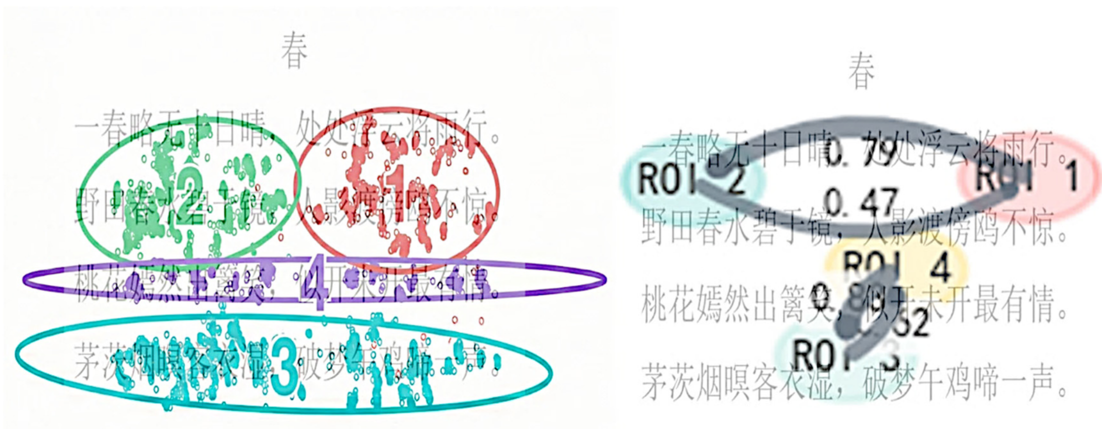

Figure 5. Visualization of the reading strategy of a representative expert (subject 5) modeled by EMHMM. The **left panel** illustrates spatial distribution of fixations (dots). Colored ellipses represent the estimated hidden states (ROIs) reflecting the reader's cognitive focus. Distinct colors denote different ROIs: Red (ROI 1), Green (ROI 2), Cyan (ROI 3), and Purple (ROI 4). The **right panel** illustrates transition network: Arrows indicate the transition probabilities between ROIs. To visualize the scanning trajectory independent of fixation duration, probabilities are normalized given gaze departure (i.e., $P(ROI_j|ROI_i)$). Thicker arrows denote higher transition probabilities.

Spatially (Figure 5, left panel), the ROIs identified for experts were significantly fewer in number but spatially more expansive, characterized by larger covariance matrices. Notably, these expert ROIs often bridged the semantic boundaries between the first half (scene description) and the second half (emotional response) of a couplet. This spatial expansion indicates that experts were engaged in semantic integration rather than mere decoding. By grouping spatially distinct but semantically related keywords—such as associating the imagery of "spring water" (Line 2) with "peach blossoms" (Line 3) into unified cognitive states—experts demonstrated a capability to process the poem as a coherent artistic entity.

Temporally (Figure 5, right panel), the transition network reveals a sophisticated "global scanning" trajectory. The expert scan path is defined by long-range, bidirectional loops between the beginning (ROI 1 & 2) and the end (ROI 3 & 4) of the poem. This pattern suggests an active, schema-driven process, where the reader continuously updates their understanding by cycling between earlier imagery and later emotional resolutions, rather than following a unidirectional timeline.

4.5. Qualitative Visualization of Cognitive Dynamics: A Transition Matrix Case Analysis

While spatial features describe where participants looked, the transition probabilities of the HMM reveal how their attention shifted dynamically over time. To capture these temporal dynamics, we visualized the switching behavior between cognitive states using Transition Probability Matrices and derived network metrics (see Figure 6). The transition matrix $A = \{a_{ij}\}$ serves as a quantitative fingerprint of the reading strategy, where each element represents the probability of shifting attention from ROI i to ROI j.

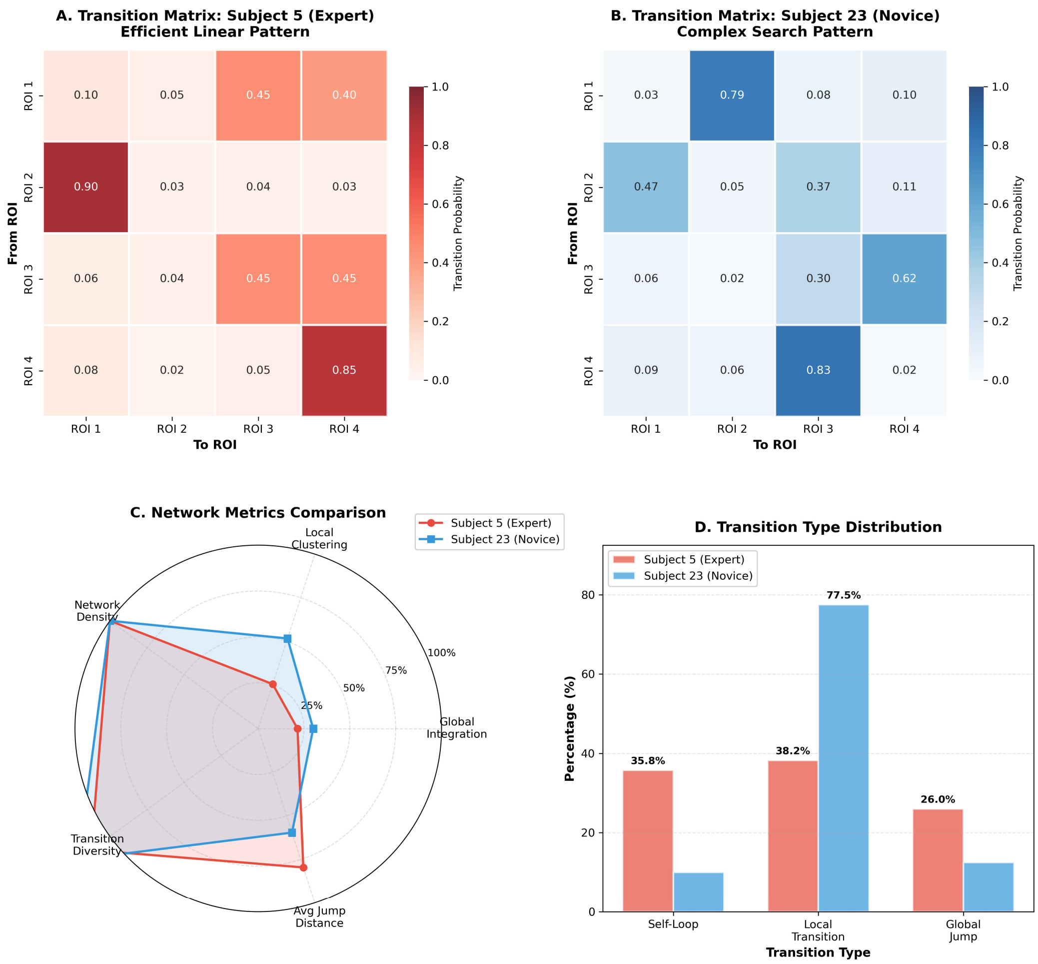

Figure 6. Case study analysis of transition patterns comparing subject 5 (expert) and subject 23 (novice). (**A**,**B**) Transition Probability Matrices: Heatmaps representing the probability of shifting attention from ROI i (y-axis) to ROI j (x-axis). Darker colors indicate higher probabilities. (**A**) Subject 5 (expert) matrix. (**B**) Subject 23 (novice) matrix. (**C**) Network Metrics Radar Chart: Comparison of four key graph-theoretical metrics: network density, local clustering, global integration, and transition diversity. red represents the expert; blue represents the novice. (**D**) Transition Type Distribution: Bar chart quantifying the proportion of self-loops, local transitions, and global jumps for each subject.

A comparative analysis of the matrices (Figure 6A,B) reveals a critical dissociation between the prototypical strategies. The novice case (Figure 6B) exhibited a pattern dominated by high probabilities along the diagonal (high a_{ii} values). This "entrapment" effect indicates that once a novice's attention enters a specific ROI, it is highly likely to remain there for consecutive fixations, reflecting a local processing mode involving intensive foveal re-inspection to resolve lexical ambiguities. Conversely, the expert case (Figure 6A) showed distributed probabilities in the off-diagonal elements (high a_{ij} where $i \neq j$). This distribution reflects a dynamic global scanning mode, where experts frequently initiate saccades between distinct semantic regions—such as shifting from "imagery" to "sentiment"—thereby creating a connected and efficient network of attention.

Further quantification via network metrics reinforces this distinction. the radar chart (Figure 6C) demonstrates that experts score higher on global integration and transition diversity, whereas novices show higher local clustering, indicative of cognitive redundancy. additionally, the transition type distribution (Figure 6D) provides precise evidence of efficiency: the expert exhibits a significantly lower percentage of self-loops (10.0%) compared to the novice (35.8%), while maintaining a higher proportion of global jumps (26.0%). these metrics collectively confirm that expert readers minimize redundant processing while maximizing information integration across the poem.

5. Discussion

The present study aimed to unveil the cognitive mechanisms underlying Classical Chinese Poetry reading by contrasting the spatiotemporal eye-movement patterns of experts and novices. By leveraging the data-driven EMHMM approach, we identified distinct reading strategies that transcend simple statistical differences. The following sections discuss these findings through the lenses of cognitive psychology and educational technology.

5.1. Interpreting Global vs. Local Strategies from A Schema Theory Perspective

Our qualitative EMHMM case study highlighted a fundamental dissociation in scanning strategies: the novice exemplar employed a "local, text-bound" strategy characterized by high self-loop probabilities and spatially constrained ROIs, whereas the expert exemplar utilized a "global, meaning-driven" strategy. This dichotomy can be effectively interpreted through Schema Theory (Anderson, 2018).

Experts possess well-developed domain-specific schemas—organized mental structures of poetic imagery and rhetorical rules—that allow them to process information in a top-down manner (Hattan et al., 2024; Smith et al., 2021). As observed in the expert exemplar's transition matrix, the reader's gaze frequently skipped across linear text to connect semantically related imagery (e.g., linking "spring water" to "peach blossoms"). This behavior aligns with recent neuroscientific findings that top-down processing dominates visual attention in proficient readers (Flower & Pylkkänen, 2024; Bandiera et al., 2025). It suggests that experts were actively testing hypotheses generated by their internal schemas, rather than passively decoding the text. They allocated cognitive resources to schema instantiation—integrating disparate textual cues into a coherent aesthetic experience, a hallmark of holistic processing observed in visual expertise (Ivy et al., 2023).

In contrast, novices lacked these sophisticated schemas. Consequently, they were forced to rely on bottom-up processing, allocating attention sequentially to individual characters to construct meaning from scratch (Breadmore & Carroll, 2018). This intensive "local processing" imposes a heavy load on working memory, as evidenced by their high fixation counts and regression rates. According to Cognitive Load Theory, when learners are overwhelmed by the intrinsic load of element interactivity (i.e., decoding difficult vocabulary and syntax), they have little remaining cognitive capacity for the germane load required for high-level aesthetic appreciation (Sepp et al., 2019; Chen et al., 2018). The EMHMM visualizations thus provide physical evidence of this cognitive bottleneck.

5.2. The Cognitive Mechanism of "Poetry Eye" (Word Refinement)

A specific focus of this study was the processing of the "Poetry Eye" (the expressive word "Laugh"/"笑"), which serves as the aesthetic pivot of the poem (Tang, 2020). Our statistical results showed no significant difference in early processing measures (First Fixation Duration) but significant differences in late processing measures (Total Fixation Count).

This temporal separation illuminates the cognitive mechanism of "word refinement". The lack of difference in the early stage indicates that lexical access—the initial retrieval of the word's literal meaning—is an automated process for both experts and novices; the word "Laugh" is not visually complex or rare. However, the significant divergence in the late stage suggests that the difficulty lies not in recognition, but in integration, consistent with the Construction-Integration Model (Kintsch, 2018).

For experts, the "Poetry Eye" functions as a retrieval cue that activates a broader emotional network. Once recognized, as suggested by our EMHMM dynamic tracking, proficient readers likely move quickly to other parts of the poem to verify the emotional context, resulting in the statistically fewer regressions to the word itself. For novices, however, the word "Laugh" created a "cognitive snag." As Lei and He (2025) note, the generation of mental imagery is critical for poetic comprehension; novices likely struggled to reconcile the personification of a flower "laughing" with the surrounding context. The increased regressions represent their repeated attempts to resolve this semantic incongruity and integrate the word into a plausible mental model (Chen et al., 2023). Thus, the 'Poetry Eye' acts as a cognitive bottleneck where the hidden disparity between expert schema instantiation and novice element interactivity is most visibly magnified.

5.3. Implications for AI-Driven Poetry Education

Beyond theoretical contributions, our findings hold profound practical implications for the architecture of next-generation intelligent tutoring systems (ITS) in the era of AI. While traditional computer-assisted language learning largely relies on outcome-based assessment, this study demonstrates that process-based assessment via eye-tracking offers more granular diagnostics (Jankowska et al., 2018; Grub et al., 2020). We suggest that the specific HMM parameters identified herein—particularly the diagonal-heavy transition matrices observed in the novice case—have the potential to serve as prototypical computational biomarkers for real-time cognitive diagnosis. This validates the feasibility of detecting learning status dynamically, expanding the utility of EMHMM beyond research into practical cognitive strategy modeling (Wang & Zhan, 2025; Hsiao et al., 2021).

Building on this diagnostic capability, future AI-driven platforms can incorporate gaze-contingent scaffolding to foster expert-like behaviors (Azevedo et al., 2022). For instance, by continuously monitoring the learner's scan path, the system can detect prolonged "novice/local" states indicative of cognitive overload (D'Mello & Graesser, 2023); in response, rather than passively displaying translations, the ITS could dynamically trigger visual cues—such as highlighting semantic connections—to guide the learner's gaze from local details to global imagery, effectively "simulating" the expert's aesthetic scanning path (Chettaoui et al., 2023). For example, the system could employ foveal-contingent masking to momentarily suppress irrelevant local details, thereby subtly 'nudging' the novice's gaze toward the long-range semantic associations characteristic of experts. Ultimately, this approach signifies a paradigm shift in poetry education from "teaching knowledge" to "training cognitive strategies." Looking forward, the integration of Large Language Models (LLMs) with such physiological data promises to create truly empathetic AI tutors capable of not only tracking attention but also interpreting the intent behind the gaze, thereby deepening the intersection of artificial intelligence and human aesthetic experience.

5.4. Limitations and Future Directions

Before detailing specific methodological constraints, it is important to contextualize this work as a foundational pilot experiment. The extended timeline of this research reflects the substantial complexity involved in the subsequent theoretical exploration and the rigorous development of the dynamic EMHMM modeling approach. Crucially, this single-modality investigation essentially establishes an early theoretical prototype for our broader, ongoing research into multimodal learning analytics, neurophysiological-driven adaptive learning systems, and the psychological mechanisms underlying human aesthetic experiences in Eastern culture.

Despite the promising findings, several limitations of this study must be acknowledged. First, regarding the methodology, while behavioral statistical analyses were conducted on the full valid sample ($N = 35$) to ensure statistical power, the dynamic EMHMM visualization employed a representative sampling strategy (selecting typical subjects from each group). While this approach effectively highlighted prototypical "global vs. local" strategies, it may not capture the full spectrum of individual variability within groups. Future research involving larger datasets could employ automated clustering algorithms (e.g., Co-clustering HMM) to classify all participants' scan paths without manual selection. Second, to strictly control experimental variables, this study utilized a single classical poem as the stimulus. Consequently, the generalizability of the findings to other poetic genres (e.g., Song Ci) or longer texts remains to be verified. Third, the participants were exclusively university students. Future studies should extend this paradigm to K-12 students to validate the developmental trajectory of these reading strategies. Finally, while eye-tracking reveals where attention is focused, it cannot fully explain why. Future work should consider integrating multimodal data, such as EEG or fNIRS, to provide a more comprehensive account of the neural mechanisms underlying aesthetic appreciation.

Ultimately, the interdisciplinary framework established in this study—bridging machine learning with the cognitive psychology of Eastern aesthetics—serves as a springboard for broader, cross-domain inquiries. Moving forward, we aim to extend this paradigm beyond poetry comprehension to encompass computational aesthetics and AI-driven art generation. By integrating comprehensive multimodal physiological signals (e.g., EEG, fNIRS, and EDA) with advanced generative models and Theory of Mind (ToM) architectures, future research will seek to decode the deeper neural mechanisms of aesthetic resonance and emotional well-being. This trajectory not only promises to advance neurophysiological-driven adaptive learning systems but also paves the way for creating empathetic AI capable of authentically understanding and engaging with humanistic cultural experiences.

6. Conclusions

This study bridges the gap between cognitive poetics and educational technology by employing eye-movement analysis with Hidden Markov Models (EMHMM) to visualize the internal cognitive dynamics of Classical Chinese Poetry appreciation. Our mixed-methods findings lend empirical support to the "global vs. local" processing hypothesis, revealing through prototypical visualization that an expert utilizes a schema-driven, global scanning strategy for semantic integration, whereas novices rely on a text-bound, local decoding strategy due to cognitive overload. Furthermore, the dissociation in processing the "Poetry Eye" highlights that the expert-novice gap lies in aesthetic integration rather than lexical recognition. Critically, these identified spatiotemporal patterns serve as potential computational biomarkers, validating the potential for AI-driven intelligent tutoring systems (ITS) to shift from static outcome assessment to dynamic, process-based interventions, such as gaze-contingent scaffolding. While limited by sample size and stimulus variety, this research lays the groundwork for future multimodal studies to further explore the neural mechanisms of aesthetic resonance and enhance automated poetry education.

Author Contributions

Conceptualization, F.X. and A.M.; methodology, F.X. and A.M.; software, F.X.; validation, F.X.; formal analysis, F.X. and A.M.; investigation, A.M. and Y.C.; resources, H.Y.; data curation, F.X. and A.M.; writing—original draft preparation, F.X.; writing—review and editing, A.M. and M.L.; visualization, F.X.; supervision, H.Y.; project administration, H.Y.; funding acquisition, H.Y. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the Center for Language Education and Cooperation, Ministry of Education, Online Course Construction Project (Grant No. SKZ2025-0184) and the Central China Normal University 2025 "Excellent Seed Cultivation" Project (Grant No. CCNU25ZZ040).

Institutional Review Board Statement

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board and the Ethics Committee of the School of Psychology at Central China Normal University.

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

The raw oculomotor dataset analyzed in the current study, along with the fully anonymized aggregated transition matrices and code snippets necessary to support the findings, are available from the corresponding author upon reasonable request.

Acknowledgments

The authors would like to thank all the students who participated in this study for their time and cooperation. We also appreciate the technical support provided by the School of Psychology at Central China Normal University. Furthermore, a preliminary version of this work was presented at the 2023 International Conference on Intelligent Education and Intelligent Research (IEIR) and published in its proceedings (Xiong et al., 2023). The current manuscript represents a substantial and independent expansion of that initial study.

Conflicts of Interest

The authors declare no conflict of interest.

References

Anderson, R. C. (2018). Role of the reader's schema in comprehension, learning, and memory. In *Theoretical models and processes of literacy* (pp. 136–145). Routledge.

Aryadoust, V. (2019). An integrated cognitive theory of comprehension. *International Journal of Listening*, *33*(2), 71–100.

Aryani, A. (2018). *Affective iconicity in language and poetry: A neurocognitive approach* [Doctoral dissertation, Freie Universität Berlin].

Azevedo, R., Bouchet, F., Duffy, M., Harley, J., Taub, M., Trevors, G., Cloude, E.; Dever, D.; Wiedbusch, M.; Wortha, F., & Cerezo, R. (2022). Lessons learned and future directions of MetaTutor: Leveraging multichannel data to scaffold self-regulated learning with an intelligent tutoring system. *Frontiers in Psychology*, *13*, 813632.

Bandiera, V., Primativo, S., Daini, R., Martelli, M., & Arduino, L. (2025). The role of top-down and bottom-up factors in parafoveal reading. *Frontiers in Cognition*, *4*, 1715617.

Breadmore, H. L., & Carroll, J. M. (2018). Sublexical and syntactic processing during reading: Evidence from eye movements of typically developing and dyslexic readers. *Journal of Cognitive Psychology*, *30*(2), 177–197.

Chen, L., Xu, X., & Lv, H. (2023). How literary text reading is influenced by narrative voice and focalization: Evidence from eye movements. *Discourse Processes*, *60*(10), 675–694.

Chen, O., Castro-Alonso, J. C., Paas, F., & Sweller, J. (2018). Extending cognitive load theory to incorporate working memory resource depletion: Evidence from the spacing effect. *Educational Psychology Review*, *30*(2), 483–501.

Chettaoui, N., Atia, A., & Bouhlel, M. S. (2023). Student performance prediction with eye-gaze data in embodied educational context. *Education and Information Technologies*, *28*(1), 833–855.

Chuk, T., Chan, A. B., Shimojo, S., & Hsiao, J. H. (2020). Eye movement analysis with switching hidden Markov models. *Behavior Research Methods*, *52*(3), 1026–1043.

Chuk, T., Crookes, K., Hayward, W. G., Chan, A. B., & Hsiao, J. H. (2017). Hidden Markov model analysis reveals the advantage of analytic eye movement patterns in face recognition across cultures. *Cognition*, *169*, 102–117.

Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavior Research Methods*, *44*(4), 1079–1100.

D'Mello, S. K., & Graesser, A. (2023). Intelligent tutoring systems: How computers achieve learning gains that rival human tutors. In *Handbook of educational psychology* (pp. 603–629). Routledge.

Drusch, G., Bastien, J. C., & Paris, S. (2014). Analysing eye-tracking data: From scan paths and heatmaps to the dynamic visualisation of areas of interest. *Advances in Science, Technology, Higher Education and Society in the Conceptual Age: STHESCA*, *20*(205), 25.

Eraslan, S., Yesilada, Y., & Harper, S. (2015). Eye tracking scanpath analysis techniques on web pages: A survey, evaluation and comparison. *Journal of Eye Movement Research*, *9*(1), 2.

Flower, N., & Pylkkänen, L. (2024). The Spatiotemporal Dynamics of Bottom–Up and Top–Down Processing during At-a-Glance Reading. *Journal of Neuroscience*, *44*(48), e0374242024.

French, R. M., Glady, Y., & Thibaut, J. P. (2017). An evaluation of scanpath-comparison and machine-learning classification algorithms used to study the dynamics of analogy making. *Behavior Research Methods*, *49*(4), 1291–1302.

Greenberg, K., & Zheng, R. (2022). Cognitive load theory and its measurement: A study of secondary tasks in relation to working memory. *Journal of Cognitive Psychology*, *34*(4), 497–515.

Grub, A. S., Biermann, A., & Brünken, R. (2020). Process-based measurement of professional vision of (prospective) teachers in the field of classroom management: A systematic review. *Journal for Educational Research Online*, *12*(3), 75–102.

Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 172.

Hattan, C., Alexander, P. A., & Lupo, S. M. (2024). Leveraging what students know to make sense of texts: What the research says about prior knowledge activation. *Review of Educational Research*, *94*(1), 73–111.

Hsiao, J. H., & Chan, A. B. (2019, July 24–27). *EMHMM: Eye Movement Analysis with Hidden Markov Models and Its Applications in Cognitive Research* [Conference session]. Annual Meeting of the Cognitive Science Society, Montreal, QC, Canada.

Hsiao, J. H., Lan, H., Zheng, Y., & Chan, A. B. (2021). Eye movement analysis with hidden Markov models (EMHMM) with co-clustering. *Behavior Research Methods*, *53*(6), 2473–2486.

Ivy, S., Rohovit, T., Stefanucci, J., Stokes, D., Mills, M., & Drew, T. (2023). Visual expertise is more than meets the eye: An examination of holistic visual processing in radiologists and architects. *Journal of Medical Imaging*, *10*(1), 015501.

Jacobs, A. M. (2018). (Neuro-) Cognitive poetics and computational stylistics. *Scientific Study of Literature*, *8*(1), 165–208.

Jankowska, D. M., Czerwonka, M., Lebuda, I., & Karwowski, M. (2018). Exploring the creative process: Integrating psychometric and eye-tracking approaches. *Frontiers in Psychology*, *9*, 1931.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329.

Kintsch, W. (2018). Revisiting the construction—Integration model of text comprehension and its implications for instruction. In *Theoretical models and processes of literacy* (pp. 178–203). Routledge.

Kui, X., Liu, N., Liu, Q., Liu, J., Zeng, X., & Zhang, C. (2022). A survey of visual analytics techniques for online education. *Visual Informatics*, *6*(4), 67–77.

Lee, H. H., Chen, Z. L., Yeh, S. L., Hsiao, J. H., & Wu, A. Y. (2021). When eyes wander around: Mind-wandering as revealed by eye movement analysis with hidden Markov models. *Sensors*, *21*(22), 7569.

Lee, M., Desy, J., Tonelli, A. C., Walsh, M. H., & Ma, I. W. (2023). The association of attentional foci and image interpretation accuracy in novices interpreting lung ultrasound images: An eye-tracking study. *The Ultrasound Journal*, *15*(1), 36.

Lei, Y., & He, X. (2025). The effect of imagery on the comprehension and aesthetic appreciation of Chinese ancient poetry. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000761

Magyari, L., Mangen, A., Kuzmičová, A., Jacobs, A. M., & Lüdtke, J. (2020). Eye movements and mental imagery during reading of literary texts with different narrative styles. *Journal of Eye Movement Research*, *13*(3), 10–16910.

Mak, M., & Willems, R. M. (2019). Mental simulation during literary reading: Individual differences revealed with eye-tracking. *Language, Cognition and Neuroscience*, *34*(4), 511–535.

Pelowski, M., Leder, H., Mitschke, V., Specker, E., Gerger, G., Tinio, P. P., Vaporova, E.; Bieg, T., & Husslein-Arco, A. (2018). Capturing aesthetic experiences with installation art: An empirical assessment of emotion, evaluations, and mobile eye tracking in Olafur Eliasson's "Baroque, Baroque!". *Frontiers in Psychology*, *9*, 1255.

Puma, S., & Tricot, A. (2019). Cognitive load theory and working memory models: Comings and goings. In *Advances in cognitive load theory* (pp. 41–52). Routledge.

Raković, M., Gašević, D., Hassan, S. U., Ruipérez Valiente, J. A., Aljohani, N., & Milligan, S. (2023). Learning analytics and assessment: Emerging research trends, promises and future opportunities. *British Journal of Educational Technology*.

Raschke, M., Blascheck, T., & Burch, M. (2013). Visual analysis of eye tracking data. In *Handbook of human centric visualization* (pp. 391–409). Springer.

Sadoski, M. (2018). Reading comprehension is embodied: Theoretical and practical considerations. *Educational Psychology Review*, *30*(2), 331–349.

Sepp, S., Howard, S. J., Tindall-Ford, S., Agostinho, S., & Paas, F. (2019). Cognitive load theory and human movement: Towards an integrated model of working memory. *Educational Psychology Review*, *31*(2), 293–317.

Smith, R., Snow, P., Serry, T., & Hammond, L. (2021). The role of background knowledge in reading comprehension: A critical review. *Reading Psychology*, *42*(3), 214–240.

Stopel, B. (2018). *From Mind to Text: Continuities and Breaks Between Cognitive, Aesthetic and Textualist Approaches to Literature* (p. 228). Taylor & Francis.

Sweller, J. (2024). Cognitive load theory and individual differences. *Learning and Individual Differences*, *110*, 102423.

Tang, J. (2020). Aesthetic psychology of the artistic conception of classical Tang poems: An analysis based on cognitive fluency. *Revista Argentina de Clínica Psicológica*, *29*(1), 824.

Wang, Y., Lu, S., & Harter, D. (2021). Towards collaborative and intelligent learning environments based on eye tracking data and learning analytics: A survey. *IEEE Access*, *9*, 137991–138002.

Wang, Z., & Zhan, P. (2025). Eye-tracking-based hidden Markov modeling for revealing within-item cognitive strategy switching. *Behavior Research Methods*, *57*(6), 175.

Xiong, F., Ma, A., & Leng, X. (2023, November 5–7). *Understanding individual differences in reading Chinese traditional poetry using hidden Markov models* [Conference session]. 2023 International Conference on Intelligent Education and Intelligent Research (IEIR), Wuhan, China. https://doi.org/10.1109/IEIR59294.2023.10391235.

Yi, T., Chang, M., Hong, S., & Lee, J. H. (2021). Use of eye-tracking in artworks to understand information needs of visitors. *International Journal of Human–Computer Interaction*, *37*(3), 220–233.

Zhang, S., Prykanowski, D. A., & Koppenhaver, D. A. (2023). Using Construction-Integration Theory to Interpret Reading Comprehension Instruction for Students with Autism Spectrum Disorder: A Systematic Review and Meta-Analysis. *Reading Research Quarterly*, *58*(1), 126–159.

Zheng, Y., Que, Y., Hu, X., & Hsiao, J. H. (2022, July 1–4). *Predicting reading performance based on eye movement analysis with hidden Markov models* [Conference session]. 2022 International Conference on Advanced Learning Technologies (ICALT) (pp. 172–176), Bucharest, Romania.