Identifying Different Thought Processes for Each Learner Using Eye-Tracking and Hidden Markov Models

Yohei Ohata[1], Satori Hachisuka[2], Kayoko Kurita[3], and Shin'ichi Warisawa[1]

[1] Department of Human and Engineered Environmental Studies, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan

[2] Applied Computer Science Course, Interfaculty Initiative in Information Studies, The University of Tokyo, Kashiwa, Chiba, Japan

[3] Center for Research and Development of Higher Education, The University of Tokyo Hongo, Tokyo, Japan

Running Head: THOUGHT PROCESS ANALYSIS

Manuscript Type: Research Article

Word Count: 4454 words (Microsoft Word Counts)

Acknowledgment

We have no conflicts of interest to disclose. This work involved human subjects in its research.

Correspondence concerning this article should be addressed to Yohei Ohata, Department of Human and Engineered Environmental Studies, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba, 277–8563. Japan. Email: yohei.ohata.ee@gmail.com

Abstract

Objective: This study proposes a method for analyzing individual learners' thought processes. **Background:** Personalized learning environments should be tailored to how learners understand and process information. To achieve this, it is crucial to develop methods that can identify differences in thought processes. **Method:** Thirteen participants were asked to generate as many criteria as possible for classifying six geometric figures into two groups. Participants' thinking styles were assessed through a questionnaire. A method combining eye-tracking data collected during the task and hidden Markov models (HMMs) was proposed to analyze the thought processes. **Results:** Participants exhibited distinct response patterns, which could be grouped into two categories: one group classified the figures based solely on their visual attributes, while the other employed more imaginative criteria that extended beyond the visual attributes. Differences in response patterns were associated with hierarchical thinking style. Although there was no significant difference in gaze fixation durations based on hierarchical thinking style scores, differences were observed in the temporal patterns of gaze transitions, which could be identified with 80% accuracy using HMMs. **Conclusion:** Learners' thought processes differences are reflected in the temporal dynamics of fixation scan path transitions and can be identified using HMMs. **Application:** The possibility of identifying thought processes suggests real-time detection of learners’ difficulties and support through adaptive interfaces tailored to their ways of understanding. The proposed method may serve as a foundational technology for future systems that analyze learners' cognitive strategies and comprehension in educational contexts.

Keywords: eye tracking, hidden Markov model, individual differences, problem-solving

Précis: This study proposes a method to analyze learners’ thought processes using eye-tracking and hidden Markov models. Experimental results show differences in thought processes

can be identified with high accuracy, suggesting potential applications for personalized learning systems.

Identifying Different Thought Processes for Each Learner Using Eye-Tracking and Hidden Markov Models

Introduction

Cultivating a wide range of skills and competencies has become an essential goal in education. Achieving this goal calls for educational practices that are tailored to the characteristics of each learner (Turner et al., 2017). However, there is currently no consensus on which learner characteristics should be prioritized, highlighting the need for deeper investigation into individual learning processes.

Within the field of learning sciences, which explores how people learn, it is widely argued that instruction should be adapted to the unique ways in which each learner constructs understanding. Understanding, in this context, is defined as the process of integrating new information into pre-existing knowledge structures and subsequently restructuring them (Davis, 1984; Sawyer, 2014a, 2014b). However, since each learner's prior experiences and existing knowledge differ, their methods of achieving understanding also vary, highlighting the importance of personalized instructional approaches (Coffield et al., 2004; Miyake, 1986; Peach et al., 2021; Shirouzu et al., 2010). Yet, quickly identifying a learner’s preferred approach to comprehension remains a significant challenge, especially in contexts where teachers and students have not yet developed rapport. In the absence of such understanding, learners may struggle to fully grasp new concepts, potentially limiting the effectiveness of instruction. Therefore, it is essential to create personalized learning environments that incorporate systems capable of identifying how each learner understands.

Establishing such environments requires robust models for analyzing the diverse cognitive approaches that learners use to make sense of information. Specifically, this involves investigating the thought processes unique to each learner. In this study, we define cognition as “the broad mental process of acquiring and understanding knowledge through perception, memory, judgment, and decision-making,” and thinking as “the conscious process of manipulating mental representations, such as ideas and images, to perform reasoning and problem-solving” (American Psychological Association, n.d.). We focus specifically on thought processes. However, because these processes are internal and not directly observable, distinguishing among them remains a major challenge. This necessitates the development of analytical models capable of inferring and differentiating thought processes in a quantifiable manner.

One widely used method for analyzing cognition is the Think Aloud protocol, in which participants verbalize their thoughts while performing a task (Ericsson & Simon, 1993). Although this method can provide valuable insights, it is limited in its ability to capture non-verbalized or unconscious cognitive activity (Cowan, 2019; Ericsson & Fox, 2011). Other methods such as neuroimaging techniques are often impractical in educational settings due to operational constraints (Janssen et al., 2021).

In recent years, eye tracking has emerged as a promising alternative for investigating cognitive processes (Eraslan et al., 2015; Lin et al., 2017; Nakano & Ishii, 2010; Salvucci & Anderson, 1999; Wang et al., 2018; Zawoyski et al., 2015). Based on the eye–mind hypothesis (Just & Carpenter, 1980; Wu & Liu, 2022), eye tracking has been applied to cognitive load assessment, learning evaluation, and educational design (Eisma, 2020; Kanaan, 2021; Kim, 2023, 2024; Shafiei, 2024). Recent studies show that gaze patterns reflect problem-solving strategies and can classify learning styles from eye-tracking data (Bittner et al., 2024; Hsing et al., 2022). Our prior work also supports the use of eye tracking for analyzing cognitive processes (Ohata et al., 2024).

These findings suggest that eye tracking is effective in visualizing learners' thought processes. However, to realize educational support tailored to individual learners' comprehension methods, it is essential to analyze thought processes based on the cognitive and thinking characteristics associated with their understanding approaches. Although most existing studies have analyzed thought processes themselves, analytical methods grounded in individual learners' cognitive and thinking characteristics have not been sufficiently established. Therefore, the purpose of this study is to propose a thought process analysis method based on cognitive and thinking characteristics. Specifically, we propose a method that analyzes thought processes by computationally modeling gaze behavior as a behavioral indicator for each learner with different cognitive and thinking characteristics. The originality of this study lies in its attempt to elucidate, from the perspective of computational models using behavioral indicators, how differences in learners' cognitive and thinking characteristics manifest as features in thought processes. This method aims to explain the relationship between learners' individual characteristics and thought processes through behavioral indicators and computational models.

Proposal of Analyzing Models for Thought Processes

We propose a thought process analysis method based on learners' cognitive and thinking characteristics. According to Kintsch & Greeno (1985) and Mayer (2013), thought processes constitute a series of cognitive processes from semantic representation of audiovisual input information to strategy planning, execution, and self-regulation. Since these processes are influenced by learners' cognitive and thinking characteristics, such as the structure of prior knowledge and tendencies in strategy selection, different thought processes are formed for each learner. Furthermore, from the perspective of the eye-mind hypothesis, such differences in thought processes manifest as variations in eye-gaze behaviors, including fixation duration and transitions. Synthesizing this information, differences in learners' cognitive and thinking characteristics generate differences in thought processes, which further manifest as differences in eye-gaze behavior.

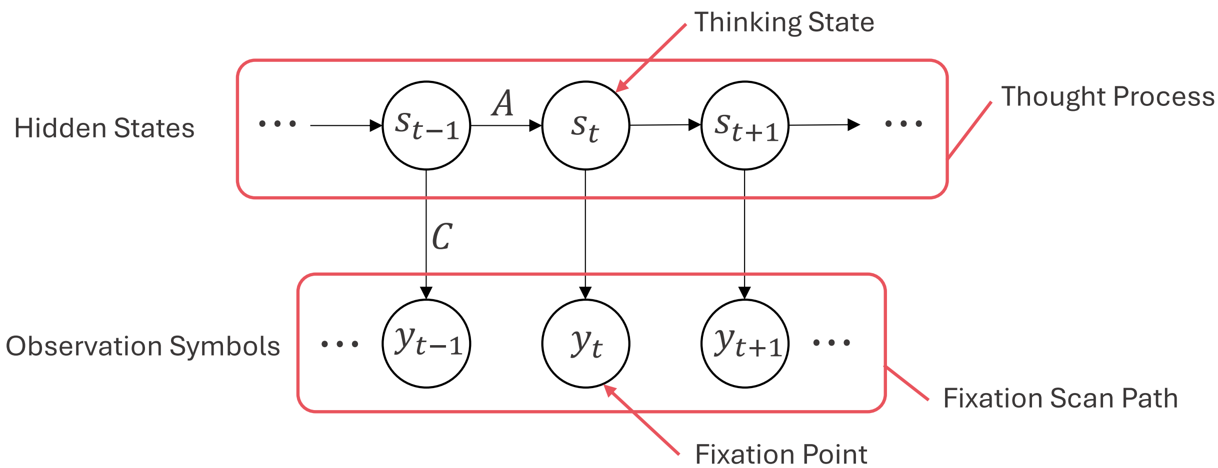
Therefore, this study proposes a thought process analysis method using hidden Markov models (HMMs), treating gaze behavior as observable states and thought processes as latent states. This method aims to visualize and analyze differences in thought processes across cognitive and thinking characteristics by constructing HMMs for each characteristic type.

The hidden Markov model is a time-series statistical model capable of probabilistically representing latent states and their dynamics underlying observable sequential data. The graphical representation of the proposed method is illustrated in Figure 1. In this study, the sequence of observation symbols $\boldsymbol{y}_{1:T} = y_1, y_2, \dots\ y_T$ represents the fixation scan path, where each y_t corresponds to a fixation point on a specific Area of Interest (AOI). The hidden states $\boldsymbol{s}_{1:T} = s_1, s_2, \dots\ s_T$ represent the underlying thought processes, with s_t denoting the latent thinking state at time t. The model further includes three parameters of probability distributions: the state transition probabilities A, which define the likelihood of transitioning from one hidden state to another; the symbol emission probabilities C, which represent the likelihood of observing a particular fixation given a hidden state; and the initial state probabilities $\boldsymbol{\pi}$, which define the probability distribution over the starting states. The hidden states s_t transition with probabilities specified in the rows of A, and at each time step emits an observation symbol y_t according to the probabilities in the rows of C. By modeling fixation scan paths using HMM, it becomes possible to analyze the underlying thought processes from observed eye-gaze behavior as dynamics of HMM's latent states.

Additionally, in this study, HMMs are constructed for each distinct cognitive and thinking characteristic. This is because modeling learners with different cognitive and thinking characteristics within a single HMM may result in averaging differences between characteristics in the state transitions and transition probabilities of the latent state space, potentially failing to adequately capture inter-individual differences in thought processes. The next chapter describes the experimental methods for verifying and evaluating whether the proposed method can analyze thought processes for individual learners.

Figure1

Graphical Representation of a Hidden Markov Model



Experimental Methods

Participants

Thirteen healthy adult volunteers (both male and female) participated in this study on a voluntary basis. Two participants were excluded from the analysis due to failure to follow task instructions, resulting in a final sample of eleven participants (mean age ± SD = 22.7 ± 6.17 years). A previous eye movement study reported a significant trend toward comparable individual differences in problem-solving strategies, using a sample size of similar magnitude to that of the present study (Fitzhugh et al., 2008). This research compiled with the tenets of the Declaration of Helsinki and was approved by the Ethics Committee of the University of Tokyo (Application No. 22-416). Informed consent was obtained from each participant.

Task

The task was to classify six different shapes (shown in Table 1) into two groups and find as many classification criteria as possible. The shapes systematically varied across four visual dimensions: color (red or blue), pattern (solid or striped), size (40 × 40 px or 200 × 200 px), and geometric form (circle, square, or triangle). Each shape was presented individually on a tablet interface (see Figure 2b). Participants were instructed to manipulate the shapes directly on the touchscreen to form groupings according to their self-defined criteria. The task continued until the participant could no longer generate new classification categories.

This task design was inspired by prior research in personality theory, specifically the "color–form problem" developed by Honkavaara (1958a, 1958b), which has been used to assess perceptual and cognitive style. The color–form problem examines how individuals judge the similarity of visual stimuli, focusing on whether they prioritize color or geometric form in their decisions. Previous findings suggest that individuals who prioritize color may exhibit higher levels of creative thinking (Honkavaara, 1958a, 1958b). Building upon this theoretical framework, the present study introduces a novel variant of the task to investigate individual differences in thought processes.

Table 1

Parameters of Figures

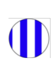
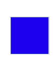
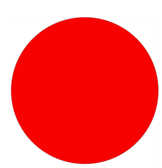
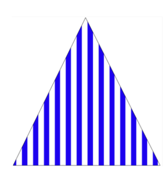
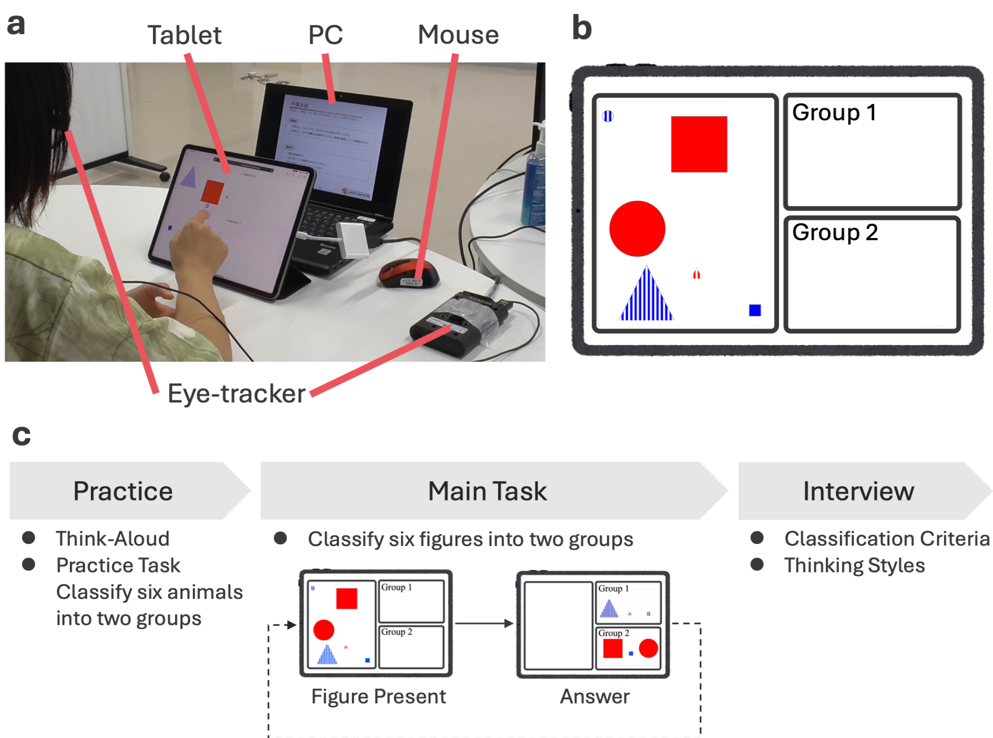
Index		1	2	3	4	5	6
Appearance							
Parameters	Shape	Circle	Triangle	Square	Circle	Triangle	Square
	Color	Blue	Red	Blue	Red	Blue	Red
	Pattern	Stripe	Stripe	Solid	Solid	Stripe	Solid
	Size	Height = 1cm, Width = 1cm			Height = 5cm, Width = 5cm		

Figure 2

Experiment. ***a*** *Experimental Setup.* ***b*** *An Example of Tablet Screen.* ***c*** *Experiment Procedure*



Investigation of Cognitive and Thinking Characteristics

Participants were also asked to complete the Thinking styles (Coffield et al., 2004; Sternberg, 1988) questionnaire in order to investigate individual differences in thought processes. Thinking styles refer to the preferred ways individuals organize, direct, and manage their thinking. This concept was proposed in Sternberg's theory of mental self-government (Sternberg, 1988), which suggests that individual thought processes are governed by the self much like a government governs a society. Using this analogy, Sternberg defined Thinking styles as five distinct categories and thirteen specific types. Unlike other psychometrics that assign a single label to a participant's cognitive or thinking profile, this approach provides a multidimensional characterization by combining several indicators of strengths and weaknesses. Moreover, prior research with university students in Hong Kong provided evidence for the cross-cultural validity of the Thinking Styles Inventory (Zhang, 1999).

The five styles and thirteen types of Thinking styles are shown in Table 2. Each item is rated on a seven-point scale, with higher scores indicating a stronger inclination toward the corresponding style. Thinking styles does not categorize participants into one of thirteen distinct styles. Instead, each participant receives a score for all thirteen styles, allowing an analysis of the balance between their strengths and weaknesses across these dimensions. In this study, we used the simplified Japanese version of the Thinking Styles Inventory (Hiruma, 2003).

Table 2

Thirteen Items of Thinking Styles

Dimension	Style	Description
Functions	Legislative	Like to create their own rules
	Execute	Like to follow rules
	Judicial	Like to evaluate rules and procedures
Forms	Monarchic	To be motivated by a single goal or need at a time
	Hierarchic	To be motivated by a hierarchy of goals, with the recognition that not all goals can be fulfilled equally well and that some goals are more important than others
	Oligarchic	To be motivated by multiple, often competing goals of equal perceived importance
	Anarchic	To be motivated by a potpourri of need
Level	Global	Prefer to deal with relatively large and abstract issues
	Local	Often like concrete problems requiring detail work
Scope	Internal	Tend to be introverted, task-oriented, aloof, socially less sensitive, and interpersonally less aware than externalists
	External	Tend to be extroverted, people oriented, outgoing, socially more sensitive, interpersonally more aware than internalist
Leaning	Conservative	Like to adhere to existing rules and procedures
	Progressive	Like to go beyond existing rules and procedures

Experimental Setup

The experimental setup is shown in Figure 2a. The task was conducted using a tablet (Apple iPad Pro 12.9 inch) placed in front of the participants. The eye movements of each participant during the task were recorded using an eye tracker (Tobii Pro Glasses 3). A PC was used to explain the experiment and answer a questionnaire about the experiment.

Procedure

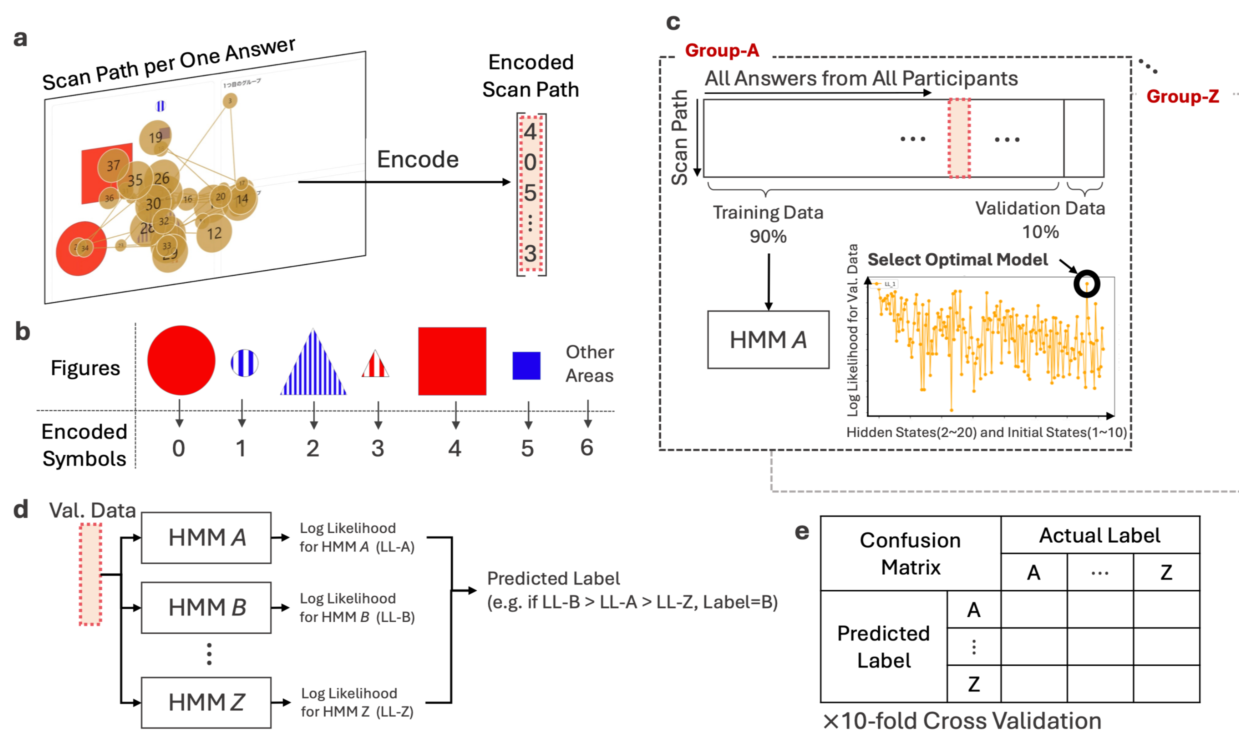
The experimental protocol comprised three sequential stages: a practice session, a main task session, and a post-task interview, as depicted in Figure 2c. During the practice session, participants engaged in a think-aloud familiarization and completed a practice task designed to resemble the main task. Initially, they were explained the think-aloud method, including its purpose and procedures, followed by a brief practice to familiarize themselves with the approach. In the practice task, participants classified six animal icons displayed on the tablet into two categories in as many distinct ways as possible. Eye-tracking data were collected during this stage to ensure proper calibration and participant comfort with the apparatus. In the main task, participants were asked to classify six geometric shapes into two groups, generating as many classification categories as possible. The shapes were displayed individually on the tablet. Participants were explicitly informed that there were no constraints on classification criteria and no time limit. Participants continued to think aloud during the task, and their eye movements were recorded. Following task completion, participants engaged in a semi-structured interview in which they were asked to describe the rationale behind each classification they had made. After the interview, participants completed a post-experiment questionnaire as well as the Thinking Styles Inventory.

Data Collection and Statistical Analysis

Four primary data types were collected in this study: task responses, thinking style scores, fixation duration data, and scan path data. The think-aloud data were used as a supplementary reference to inform the analysis. Task responses consisted of the multiple classification criteria generated by each participant in response to the figure classification task. Thinking styles were assessed using participants' scores across thirteen dimensions. To explore the relationship between thinking styles and thought processes, participants were grouped according to patterns observed in their classification responses. For each group, mean scores across the thirteen thinking styles were computed. To identify statistically significant differences between groups, two-tailed independent samples t-tests were conducted for each thinking style. The significance threshold was set at $\alpha = 0.05$. Fixation and scan path data were collected using the eye-tracking system. A fixation was defined as a continuous gaze maintained on a specific Area of Interest (AOI) for 300 milliseconds or longer, in accordance with previous research (Solso, 2003). To account for individual variability in overall fixation behavior, fixation durations were z-transformed within participants prior to statistical analysis. To investigate the relationship between gaze behavior and thinking styles, z-transformed fixation duration data were analyzed using a linear mixed-effects model (LMM). In this model, figure type and thinking style group (based on problem-solving strategy classification) were included as fixed effects, along with their interaction term. Participant was included as a random intercept to control for repeated measures and inter-individual differences. Scan path data represented the sequential transitions of gaze between AOIs, capturing the temporal dynamics of visual exploration. These data were modeled using a hidden Markov model.

Figure 3

Analyzing Method for Identification of Differences in Thought Process. **a** *Extracting Scan Path Per One Answer.* **b** *Encoding Rule for Fixation Sequences.* **c** *Training the Hidden Markov Model (HMM).* **d** *Model Inference from Validation Data.* **e** *Evaluation and Cross Validation.*



Construction of the Fixation Scan Path Dataset

Figure 3a, 3b illustrate the method used to construct the dataset. We constructed the dataset for the HMM by extracting the data $\boldsymbol{y}_{1:T}$ from fixation scan paths recorded during tasks. Figure 3a represents the fixation scan path from when a problem is presented until the participant begins moving the figures while considering classification criteria. The brown circles indicate fixation points. The size of each circle indicates the duration of the fixation; larger circles mean a longer fixation. We created the dataset by applying the encoding rule shown in Figure 3b, extracting the fixation point sequence from the overlapping areas of the brown circle and the figure. Each figure is assigned a unique numerical label, and regions outside the displayed figures are also assigned numbers.

Model Fitting and Validation Procedure

To verify the effectiveness of the proposed method, this study conducted two analyses with different objectives. The first analysis verified whether differences in thought processes could be identified as dynamics in eye-gaze behavior. This procedure is shown in Figure 3. First, a fixation scan path dataset was created from all responses of all participants (Figure 3a, b). Next, participants were grouped based on differences in Thinking Styles related to problem-solving strategies, and datasets and HMMs were constructed for each group. Each dataset was divided into training and validation sets, and HMMs were trained using the training data (Figure 3c). For HMM parameter estimation, the Variational Bayesian Expectation Maximization (VB-EM) algorithm was adopted, which suppresses local optima and overfitting while automatically adjusting model complexity (Beal, 2003; Beal & Ghahramani, 2003; Yang et al., 2015). Observation sequences were treated as categorical variables, with categorical distributions for emission probabilities and Dirichlet distributions for prior distributions. Regarding model structure, combinations of hidden state numbers (2–20) and initial state numbers (1–10) were explored through grid search, and the condition that maximized log-likelihood on validation data was adopted as the optimal model for each dataset (Figure 3c). Subsequently, validation data from all groups were input to all trained models to calculate log-likelihoods, and the model providing maximum likelihood was identified as the predicted label (Figure 3d). Based on the obtained predicted labels, accuracy, recall, and F1 scores were calculated using 10-fold cross-validation (Figure 3e).

The second analysis examined structural differences in thought processes within the HMM's latent space. Specifically, state transition matrices were extracted from HMMs trained using all data from each group, and their structures were compared. Here, to directly compare transition structures and probabilities on the same scale, the number of hidden states was unified to 7 across all groups. This decision was made not to regard hidden states as AOIs themselves, but rather, considering that information sources in this task are consolidated into seven AOIs, to align the resolution of cognitive focus—"which information source the thought is directed toward"—with the granularity of observable information units (AOIs). Furthermore, weighted directed graphs were created from the estimated transition matrices, and the PageRank algorithm (Page et al., 1999) was used to evaluate the structural centrality that each latent state plays in the dynamics of thought processes.

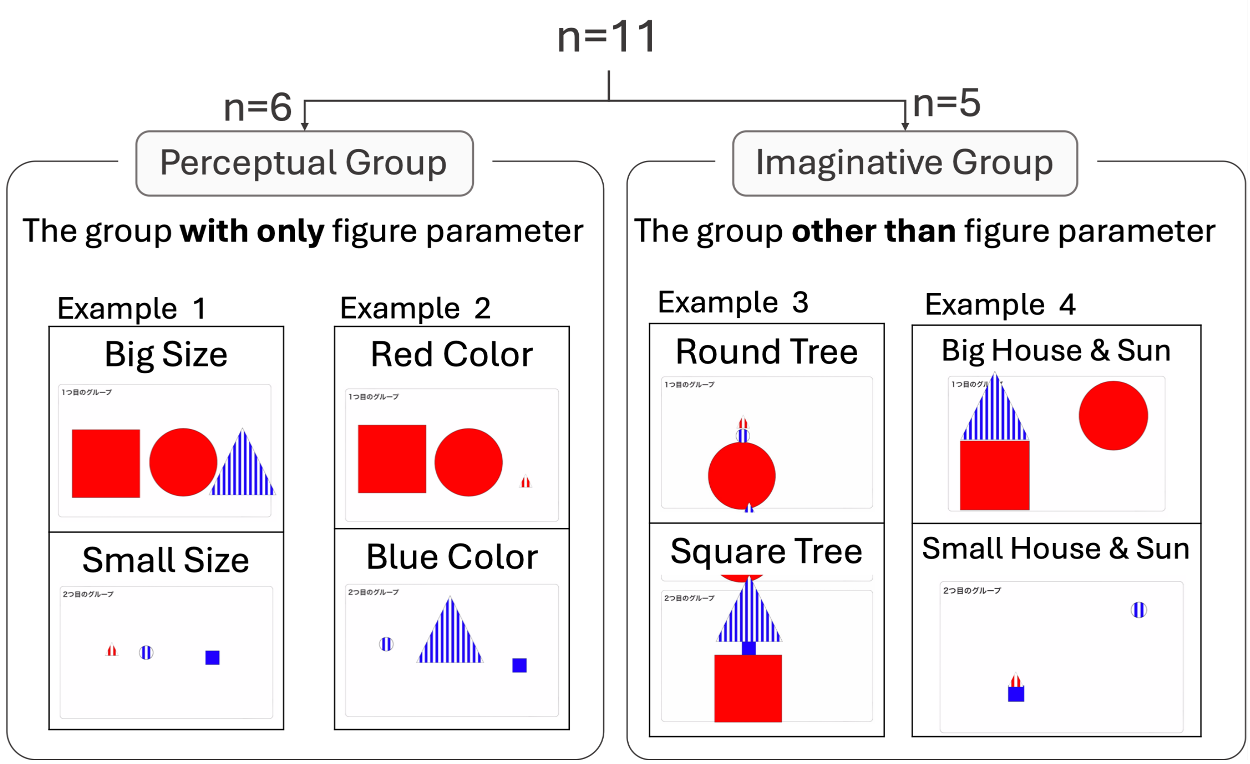
Results

Participant Response Classification

A total of 97 responses were collected from 11 participants. Based on the tendencies of their classification criteria, participants were categorized into two groups, as illustrated in Figure 4. The first group (n = 6) provided classification strategies grounded in the physical properties of the stimuli specifically, color, pattern, size, and geometric features such as corners or lines, as outlined in Table 1. These strategies included responses based on single parameters or combinations thereof. This group is hereafter referred to as the Perceptual Group. The second group (n = 5) provided answers that were not only related to the figure parameters, but also to the relatively free imagination. These responses included classifications based on combinations or relationships among figures, associations with symbolic or character-like features, visual impressions, and the spatial relationship between objects. This group is referred to as the Imaginative Group.

Figure 4

Differences in Answers to the Task

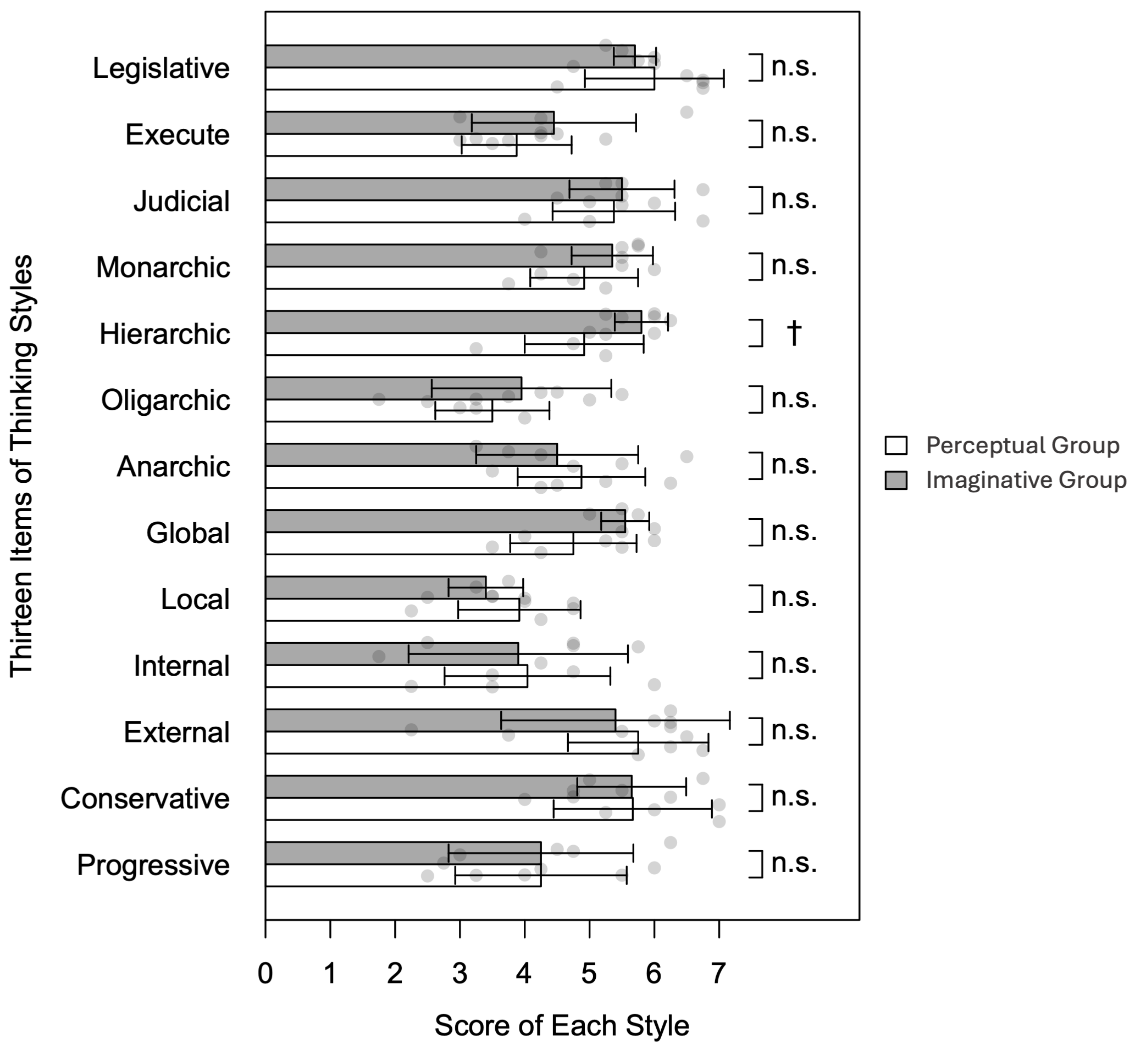


Thinking Style Comparison

To explore the relationship between individual differences in thinking style and problem-solving approaches, mean scores for 13 thinking styles were compared between the perceptual and imaginative group. As shown in Figure 5, a trend toward significance was observed for the hierarchical thinking style, with the perceptual group exhibiting lower scores than the imaginative group ($t(7) = 2.12$, $p = .07$, $d = 0.39$). No significant differences emerged for the remaining 12 styles. Hierarchical thinking style, as defined in Sternberg's theory (1988), reflects an individual's tendency to organize goals in a structured hierarchy, prioritizing them based on importance. These results demonstrate that participants in the perceptual group and imaginative group may differ in their thinking approaches, particularly in terms of hierarchical thinking.

Figure 5

Comparison of Average Scores across Groups for the Thirteen Items of Thinking Styles



Fixation Duration Analysis

To examine the effect of thinking style associated with problem-solving approach on fixation duration, we constructed a linear mixed-effects model with thinking style group (High-

Hierarchical vs. Low-Hierarchical), figure type (Index #1-6 as shown in Table1), and their interaction as fixed effects. Participant was included as a random intercept to account for inter-individual variability. Participants were divided based on their hierarchical thinking scores: those scoring above the group mean were assigned to the High-Hierarchical group (n = 6), and those below the mean to the Low-Hierarchical group (n = 5). Model results indicated no significant main effect of hierarchical thinking style on fixation duration ($\beta = -0.17, SE = 0.33, t(48) = -0.52, p = .609$). In contrast, a significant main effect of figure type was observed: participants exhibited longer fixations for the figure #5 (blue, triangle, stripe, big) ($\beta = 1.81, SE = 0.33, t(48) = 5.43, p < .001$) and the figure #6 (red, square, solid, big) ($\beta = 1.51, SE = 0.33, t(48) = 4.56, p < .001$) compared to the reference figure#1 (blue, circle, stripe, small). Other figure types did not significantly differ from the reference. No interaction between thinking style and figure type reached significance (p > .05). These results suggest that fixation duration is modulated primarily by visual characteristics of the stimuli, rather than by differences in hierarchical thinking style.

Table 4

Fixed Effects from Linear Mixed-Effects Model Predicting Fixation Duration by Thinking Style and Figure Type

Fixed Effect	**Estimate (SE)**	**t (df)**	**p**
Intercept	- 0.623 (0.235)	- 2.65 (48)	.011*
Low-Hierarchical x figure1(blue, circle, stripe, small)			
Group (High- Hierarchical Thinking Style)	- 0.171 (0.332)	- 0.52 (48)	.609
Figure Type			
2 (red, triangle, stripe, small)	- 0.100 (0.332)	- 0.30 (48)	.764
3 (blue, square, solid, small)	- 0.021 (0.332)	0.06 (48)	.951
4 (red, circle, solid, big)	0.540 (0.332)	1.62 (48)	.111
5 (blue, triangle, stripe, big)	1.806 (0.332)	5.43 (48)	<.001***
6 (red, square, solid, big)	1.514 (0.332)	4.56 (48)	<.001***
Group x Figure Type			
2 (red, triangle, stripe, small)	- 0.014 (0.470)	- 0.03 (48)	.976

3 (blue, square, solid, small)	0.354 (0.470)	0.75 (48)	.455
4 (red, circle, solid, big)	0.724 (0.470)	1.54 (48)	.130
5 (blue, triangle, stripe, big)	0.334 (0.470)	0.71 (48)	.481
6 (red, square, solid, big)	- 0.369 (0.470)	- 0.79 (48)	.436

Discriminative Performance and Structural Comparison of Thought Processes

To further investigate the effect of thinking styles on problem-solving approaches, we analyzed participants' scan paths using hidden Markov models (HMMs). Following the procedure illustrated in Figure 3, separate HMMs were trained for the High-Hierarchical and Low-Hierarchical groups using 10-fold cross-validation with optimized state numbers. The classification performance yielded an average accuracy of 0.80 ($SD = 0.12$), average recall of 0.82 ($SD = 0.14$), and average F1 score of 0.84 ($SD = 0.10$). These results may suggest that the HMMs captured and distinguished group-specific scan path dynamics, indicating the possibility that the model reflects differences in underlying thought processes.

Next, full dataset HMMs were trained for each group, and both state transition matrices and symbol emission distributions were extracted (see Figures 6a and 6b). To evaluate the structural centrality of hidden states, we generated weighted directed graphs from the transition matrices and applied the PageRank algorithm. In HMM on Low-Hierarchical group, State 0 and State 1 serve as network centers, whereas in HMM on High-Hierarchical group, State 3 and State 5 are comparatively more central. These structural differences are also evident in the graph structures displayed in Figure7, where node size reflects PageRank centrality and edge thickness corresponds to transition probabilities. These findings suggest that the temporal dynamics of scan paths are reflected in differences among the HMM hidden states and identified between groups.

Figure 6

Hidden Markov Model Transition Matrices and Symbol Output Distributions for the Low-Hierarchical and High-Hierarchical Thinking Style Groups

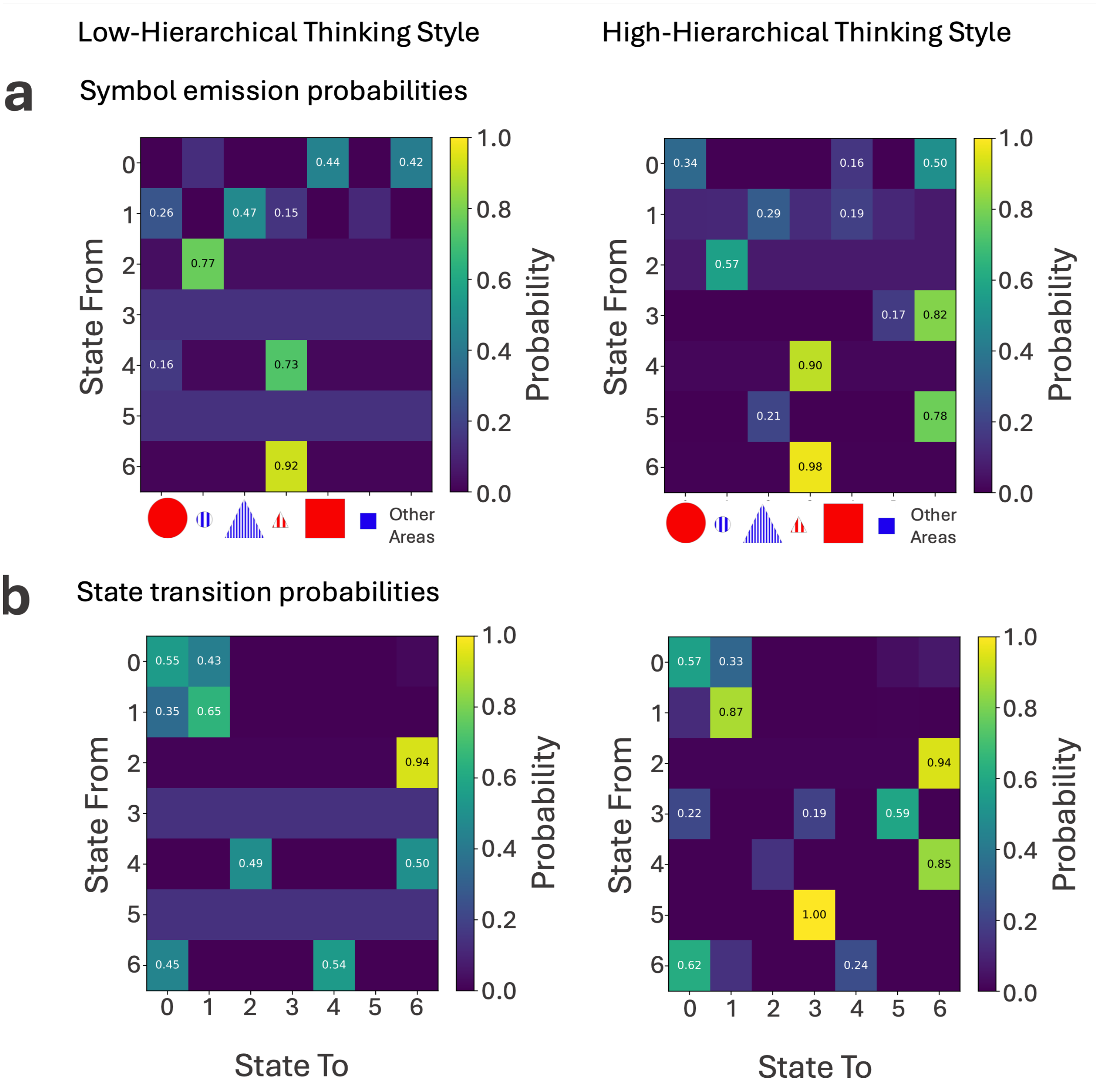
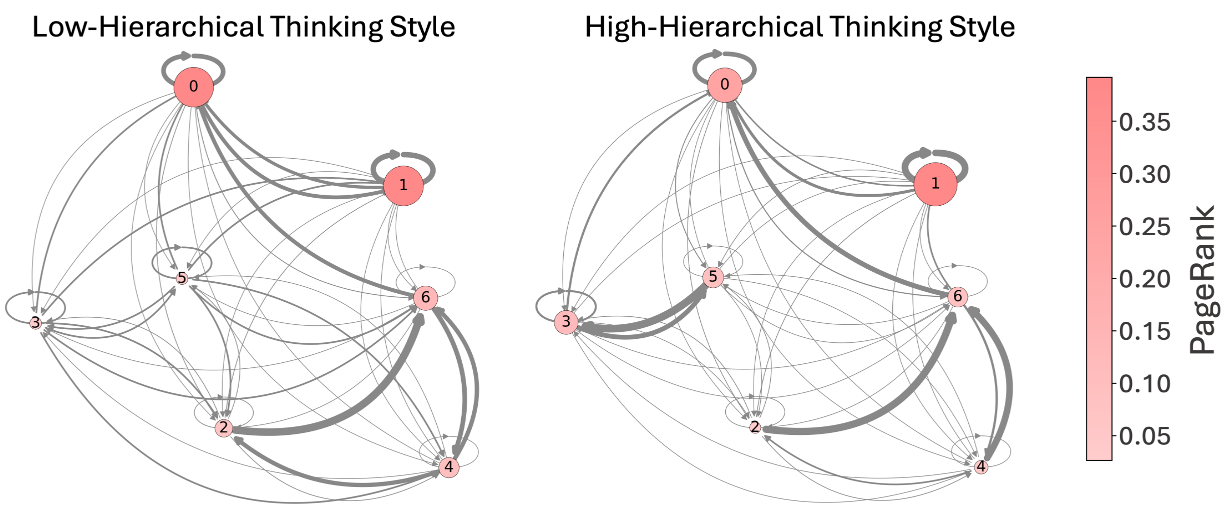


Figure 7

Graph Representations of State Transition Matrices for Low-Hierarchical and High-Hierarchical Thinking Style Groups. Node Size Reflects PageRank Centrality, and Edge Thickness Corresponds to Transition Probability



Discussion

In this study, participants repeatedly classified six different figures into two groups. As a result, their responses were categorized into two patterns: a Perceptual Group (based on the visual attributes of the figures) and an Imaginative Group (based on more creative interpretations). Participants in the imaginative group exhibited a stronger tendency toward a hierarchical thinking style. Although there were no significant differences in fixation durations on individual figures depending on thinking style, differences in scan path patterns were observed between the High-and Low-Hierarchical groups, and these patterns could be identified with 80% accuracy using HMMs.

These differences in response tendencies and thinking styles may reflect the operation of distinct thought processes among learners. According to Sternberg, "A hierarchical style refers to the tendency to solve problems by organizing them in a gradual and structured manner." Participants with a stronger hierarchical thinking style were more likely to provide creative responses. For example, they might have interpreted a triangle as the "roof of a house," suggesting that they associated figure parameters with their own knowledge and mental imagery. These findings indicate that, even when engaging in the same task, individuals may employ different thought processes depending on their thinking styles.

Although no significant group differences in fixation duration were found, fixation duration varied depending on the visual characteristics of the figures. Bottom-up attentional mechanisms driven by visual saliency (e.g., figure size, color, and pattern) strongly influenced gaze behavior. Prior research has shown that stimulus attributes such as size and surface complexity modulate fixation time (Nuthmann et al., 2016), and similar effects appear to underlie the present findings. Therefore, in this study, figure size and the presence or absence of patterns may have influenced fixation durations, suggesting that the differences might be due to visual saliency rather than differences in thought processes.

By contrast, the scan path analysis using hidden Markov models (HMMs) revealed clear distinctions between groups. The models successfully classified participants' scan paths with high accuracy (80% on average), indicating that differences in thought processes were embedded in the temporal dynamics of eye movements. This finding aligns with prior work demonstrating that scan path structures vary with task type and problem-solving strategy (Guo et al., 2024; Laurence et al., 2018). These findings suggest that scan path patterns are influenced by top-down attention driven by task demands, and that such patterns differ across groups. Thus, even with the same task, learners may engage in different thought processes. These individual differences are not captured by static indicators like fixation duration but are instead reflected in dynamic indicators such as scan paths, which can be visualized and classified using time-series statistical models such as HMMs.

This study has several limitations. First, although differences were observed not only in answers but also in cognitive and thinking characteristics through the experiment and analysis, these findings may differ depending on the age group and sample size of the participants. Therefore, at this time, it is necessary to continuously verify the trends seen in the answers and cognitive and thinking characteristics of new participants. To address this, future research should focus on developing generalized models that can predict how differences in cognitive and thinking characteristics are reflected in variance in problem-solving answers. Second, the proposed method for analyzing thought processes showed that thought processes are identifiable and the differences can be expressed by the probability distribution of HMMs, but it remains unclear what specific features characterize human thought processes. In response to this, by integrating eye movement and brain data, the correspondence between gaze patterns and neural substrates can be clarified and thought processes can be modeled.

The feature of this proposal is that it demonstrates that it is possible to "indirectly" analyze thought processes using a glasses-type eye tracking device, rather than using methods such as neuroimaging techniques that "directly" measure brain function. Analysis of thought processes in traditional learning environments has relied on externalized expressions such as notes and verbalization (Ridgway & Cox, 2024). In contrast, in recent learning environments, it is essential for teachers to understand learners' thinking patterns early on. The fact that this study suggested the possibility of identifying thought processes using scan path indicates that it may become possible to detect learners’ difficulties in real-time by further advancing the analysis of thought processes.

Furthermore, this technology may enable the development of adaptive interfaces that provide real-time support optimized to each learner’s way of understanding and thinking style, in accordance with their current thinking state. In addition, minimizing the constraints imposed on learners is expected to become one of the important system requirements in school education. The proposed method is promising as a basic technology for systems that analyze understanding methods because it reduces physical constraints compared to brain function measurement technology and is thought to not hinder learners' thinking. The introduction of such a system is expected to have the following effects: learners will be able to understand their own thinking tendencies and select a learning method that suits them, and educators will be able to teach in a way that is easier for learners to understand.

Conclusion

The purpose of this study was to propose a method for analyzing the differences and characteristics of each learner's thought process. Although all participants were asked to generate as many different criteria as possible for dividing six geometric figures into two categories, their responses varied, suggesting that different thinking styles may be associated with distinct thought processes. While fixation duration did not significantly differ between groups defined by differences in thinking styles associated with problem-solving approaches, gaze scan paths differed depending on thinking style groups and could be classified with 80% accuracy using hidden Markov models. This opens up bright prospects for the realization of an adaptive and tailored learning system that analyzes learners' unique ways of understanding.

Key Points

- Participants performed geometric figures classification task while their eye movements were recorded.
- An analyzing method for identifying different thought processes utilizing eye movement scan paths modeled with hidden Markov models was proposed.
- Differences in learners' thought processes were embedded in their eye movement patterns, enabling group classification with 80%.
- The proposed method may serve as a foundational technology for personalized learning systems that evaluate the way each learner comprehends.

References

American Psychological Association. (n.d.). *Cognition*. In *APA dictionary of psychology*. https://dictionary.apa.org/cognition

American Psychological Association. (n.d.). *Thinking*. In *APA dictionary of psychology*. https://dictionary.apa.org/thinking

Beal, M. J. (2003). Variational algorithms for approximate Bayesian inference [Doctoral dissertation, University College London].

Beal, M. J., & Ghahramani, Z. (2003). The variational Bayesian EM algorithm for incomplete data: with application to scoring graphical model structures. *Bayesian Statistics, 7*, 453-464, Oxford University Press.

Bittner, D., Nadimpalli, V. K., Grabinger, L., Ezer, T., Hauser, F., & Mottok, J. H. (2024). Uncovering learning styles through eye tracking and artificial intelligence. In *Proceedings of the 2024 ACM Symposium on Eye Tracking Research and Applications* (pp. 1–7). Association for Computing Machinery. https://doi.org/10.1145/3649902.3653940

Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning: A systematic and critical review.* Learning and Skills Research Centre. https://doi.org/10.13140/2.1.3999.4889

Cowan, J. (2019). The potential of cognitive think-aloud protocols for educational action-research. *Active Learning in Higher Education, 20*(3), 668–682.

Davis, R. B. (1984). *Learning mathematics: The cognitive science approach to mathematics education.* Croom Helm.

Eisma, Y. B., Borst, C., Paassen, R. van, & Winter, J. de. (2020). Augmented Visual Feedback: Cure or Distraction? *Human Factors, 63*(7), 1156-1168. https://doi.org/10.1177/0018720820924602

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). https://doi.org/10.16910/jemr.9.1.2

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data (Revised ed.)*. MIT Press. https://doi.org/10.7551/mitpress/5657.001.0001

Ericsson, K. A., & Fox, M. C. (2011). Thinking aloud is not a form of introspection but a qualitatively different methodology: reply to Schooler. *Psychol Bulletin, 137*(2), 351-354. https://doi.org/10.1037/a0022388

Fitzhugh, S., Shipley, T. F., Newcombe, N., McKenna, K., & Dumay, D. (2008). Mental rotation of real-word Shepard-Metzler figures: An eye tracking study. *Journal of Vision, 8*(6), 648. https://doi.org/10.1167/8.6.648

Guo, Y., Pannasch, S., Helmert, J. R., & Kaszowska, A. (2024). Ambient and focal attention during complex problem-solving: preliminary evidence from real-world eye movement data. *Frontiers in Psychology, 15*. https://doi.org/10.3389/fpsyg.2024.1217106

Hsing, H.-W., Lau, N., & Bairaktarova, D. (2022). Relationships of eye gaze metrics between cognitive processes and strategy in spatial problem-solving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 66*(1), 480-484. https://doi.org/10.1177/1071181322661125

Hiruma, T. (2003). Thinking styles and test taking behavior of university students. *Kansei University Institutional Repository, 34*, 11–37.

Honkavaara, S. (1958a). Comparison of the relation of color and form reactors at Harvard and London University. *Journal of Psychology, 46*, 23–24. https://doi.org/10.1080/00223980.1958.9916296

Honkavaara, S. (1958b). Relation of interpersonal preference and emotional attitude of the subjects. *Journal of Psychology, 46*, 25–31. https://doi.org/10.1080/00223980.1958.9916297

Janssen, T. W. P., Grammer, J. K., Bleichner, M. G., Bulgarelli, C., Davidesco, I., Dikker, S., Jasińska, K. K., Siugzdaite, R., Vassena, E., Vatakis, A., Zion-Golumbic, E., & van Atteveldt, N. (2021). Opportunities and limitations of mobile neuroimaging technologies in educational neuroscience. *Mind, Brain, and Education, 15*(4), 354–370. https://doi.org/10.1111/mbe.12302

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Kanaan, D., & Moacdieh, N. M. (2021). Eye Tracking to Evaluate the Effects of Interruptions and Workload in a Complex Task. *Human Factors, 64*(7), 1168-1180. https://doi.org/10.1177/0018720821990487

Kim, J. H., Yu, C.-Y., Seo, K., Wang, F., & Oprean, D. (2023). The Effect of Virtual Instructor and Metacognition on Workload in a Location-Based Augmented Reality Learning Environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 67*(1), 1550-1555. https://doi.org/10.1177/21695067231192938

Kim, J. H., Monhanty, S., Pulipati, V., Wang, F., Mostowfi, S., Oprean, D., Wang, Y., & Seo, K. (2024). Assessing Student Performance Through Pupil Dilation and Problem-Solving Time in Augmented Reality. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 68*(1), 589-595. https://doi.org/10.1177/10711813241261682

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*(1), 109–129. https://doi.org/10.1037/0033-295X.92.1.109

Laurence, P. G., Mecca, T. P., Serpa, A., Martin, R., & Macedo, E. C. (2018). Eye movements and cognitive strategy in a fluid Intelligence Test: item type analysis. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.00380

Lin, Y. Y., Holmqvist, K., Miyoshi, K., & Ashida, H. (2017). Effects of detailed illustrations on science learning: An eye-tracking study. *Instructional Science, 45*(5), 557–581. https://doi.org/10.1007/s11251-017-9417-1

Mayer, R. E. (2013). Problem solving. In D. Reisberg (Ed.), *The Oxford handbook of cognitive psychology* (pp. 769–778). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195376746.013.0048

Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive Science, 10*(2), 151–177. https://doi.org/10.1207/s15516709cog1002_2

Nakano, Y., & Ishii, R. (2010). Estimating user's engagement from eye-gaze behaviors in human-agent conversations. *Proceedings of the 15th International Conference on Intelligent User Interfaces* (pp. 139–148). Association for Computing Machinery. https://doi.org/10.1145/1719970.1719990

Nuthmann, A. (2016). Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task. *Psychonomic Bulletin & Review, 24*(2), 370–392. https://doi.org/10.3758/s13423-016-1124-4

Ohata, Y., Hachisuka, S., Kurita, K., & Warisawa, S. (2024). Visualizing and revealing the difference of learner's thought process and question solving method. In *Proceedings of the 12th International Conference on Information and Education Technology (ICIET '24)* (pp. 472–476). Yamaguchi, Japan.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web. *The Web Conference, 98*, 161–172.

Peach, R. L., Greenbury, S. F., Johnston, I. G., Yaliraki, S. N., Lefevre, D. J., & Barahona, M. (2021). Understanding learner behaviour in online courses with Bayesian modelling and time series characterization. *Scientific Reports, 11*, 2823. https://doi.org/10.1038/s41598-021-81709-3

Prytz, E. G., Norén, C., & Jonson, C.-O. (2018). Fixation Differences in Visual Search of Accident Scenes by Novices and Expert Emergency Responders. *Human Factors, 60*(8), 1219-1227. https://doi.org/10.1177/0018720818788142

Ridgway, L. M., & Cox, T. (2024). Investigating student approaches to rearranging circuit diagrams. *IEEE Transactions on Education*, *67*(5), 681–688. https://doi.org/10.1109/te.2024.3410375

Salvucci, D. D., & Anderson, J. R. (1999). Tracing eye movement protocols with cognitive process models. *Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp. 923–928). Erlbaum. https://doi.org/10.1184/R1/6618869.v1

Sawyer, R. K. (2014a). *The Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 1-20). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526

Sawyer, R. K. (2014b). Project-Based Learning. In Joseph, S., K., & Namsoo, S. (Eds.), *The Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 275-297). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526

Shafiei, S. B., Shadpour, S., & Mohler, J. L. (2024). An Integrated Electroencephalography and Eye-Tracking Analysis Using eXtreme Gradient Boosting for Mental Workload Evaluation in Surgery. *Human Factors, 0*(0). https://doi.org/10.1177/00187208241285513

Shirouzu, H., Miyake, N., & Masukawa, H. (2010). Cognitively active externalization for situated reflection. *Cognitive Science, 26*(4), 469–501. https://doi.org/10.1016/S0364-0213(02)00066-6

Solso, R. L. (2003). *The psychology of art and the evolution of the conscious brain*. MIT Press. https://doi.org/10.7551/mitpress/5673.001.0001

Sternberg, R. J. (1988). Mental self-government: A theory of intellectual styles and their development. *Human Development, 31*(4), 197–224. https://doi.org/10.1159/000275810

Turner, W.D., Solis, O.J., & Kincade, D.H. (2017). Differentiating Instruction for Large Classes in Higher Education. *International Journal of Teaching and Learning in Higher Education, 29*(3), 490-500.

Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning: Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology, 110*(2), 250–268. https://doi.org/10.1037/edu0000212

Wu, C. J., & Liu, C.-Y. (2022). Refined use of the eye-mind hypothesis for scientific argumentation using multiple representations. *Instructional Science, 50*(4), 551–569. https://doi.org/10.1007/s11251-022-09581-w

Yang, F., Balakrishnan, S., & Wainwright, M. J. (2015). Statistical and computational guarantees for the Baum-Welch algorithm. *arXiv preprint arXiv:1512.08269*. https://doi.org/10.48550/arXiv.1512.08269

Zawoyski, A. M., Ardoin, S. P., & Binder, K. S. (2015). Using eye tracking to observe differential effects of repeated readings for second-grade students as a function of achievement level. *Reading Research Quarterly, 50*(2), 171–184. https://doi.org/10.1002/rrq.92

Zhang, L. F. (1999). Further cross-cultural validation of the theory of mental self-government. *Journal of Psychology, 133*(2), 165–181. https://doi.org/10.1080/00223989909599731

Biographies

Yohei Ohata is currently with Center for Information and Neural Networks, National Institute of Information and Communications Technology. He received his M.S. in Human Engineered Environmental Studies from The University of Tokyo in 2024.

Satori Hachisuka is currently a lecturer with Interfaculty Initiative in Information Studies, The University of Tokyo. She received her Ph.D. in Human Engineered Environmental Studies from The University of Tokyo in 2007.

Kayoko Kurita is currently a professor with Center for Research and Development of Higher Education, The University of Tokyo. She received her Ph.D. in Education from The University of Tokyo in 2002.

Shin'ichi Warisawa is currently a professor with Graduate School of Frontier Sciences, The University of Tokyo. He received his Ph.D. in Engineering from The University of Tokyo in 1994.