Eye movement analysis with switching hidden Markov models

Tim Chuk¹ · Antoni B. Chan² · Shinsuke Shimojo³ · Janet H. Hsiao^{1,4}

Published online: 11 November 2019
© The Psychonomic Society, Inc. 2019

Abstract

Here we propose the eye movement analysis with switching hidden Markov model (EMSHMM) approach to analyzing eye movement data in cognitive tasks involving cognitive state changes. We used a switching hidden Markov model (SHMM) to capture a participant's cognitive state transitions during the task, with eye movement patterns during each cognitive state being summarized using a regular HMM. We applied EMSHMM to a face preference decision-making task with two pre-assumed cognitive states—exploration and preference-biased periods—and we discovered two common eye movement patterns through clustering the cognitive state transitions. One pattern showed both a later transition from the exploration to the preference-biased cognitive state and a stronger tendency to look at the preferred stimulus at the end, and was associated with higher decision inference accuracy at the end; the other pattern entered the preference-biased cognitive state earlier, leading to earlier above-chance inference accuracy in a trial but lower inference accuracy at the end. This finding was not revealed by any other method. As compared with our previous HMM method, which assumes no cognitive state change (i.e., EMHMM), EMSHMM captured eye movement behavior in the task better, resulting in higher decision inference accuracy. Thus, EMSHMM reveals and provides quantitative measures of individual differences in cognitive behavior/style, making a significant impact on the use of eyetracking to study cognitive behavior across disciplines.

Keywords Hidden Markov model · Eye movement · Preference decision making · EMHMM

Recent research has shown that people have idiosyncratic eye movement patterns in visual tasks that are consistent across different stimuli and tasks (e.g., Andrews & Coppola, 1999; Castelano & Henderson, 2008; Kanan, Bseiso, Ray, Hsiao, & Cottrell, 2015; Poynter, Barber, Inman, & Wiggins, 2013). These idiosyncratic eye movement patterns may reflect individual differences in cognitive style or abilities. For example, Risko, Anderson, Lanthier, and Kingstone (2012) found that participants who demonstrated higher levels of curiosity made significantly more fixations in a scene-viewing task than did

those who demonstrated lower levels of curiosity. Wu, Bischof, Anderson, Jakobsen, and Kingstone (2014) found that when viewing human faces, those who scored higher on extraversion and agreeableness personality traits looked at the eyes significantly more often than did those who scored lower. Sekiguchi (2011) found that people with good and with bad face memory performance showed different eye movement patterns during face learning.

To better understand the association between eye movement patterns in visual tasks and individual differences in cognitive style or abilities, we recently developed a state-of-the-art eye movement data analysis method, eye movement analysis with hidden Markov models (EMHMM), which takes individual differences in the temporal and spatial dimensions of eye movements into account (Chuk, Chan, & Hsiao, 2014; an EMHMM Matlab toolbox is available at <http://visal.cs.cityu.edu.hk/research/emhmm/>). The EMHMM method is based on the assumption that during a visual task, the currently fixated region of interest (ROI) depends on the previously fixated ROI. Thus, eye movements in a visual task may be considered a Markovian stochastic process, which can be better understood using a hidden Markov model (HMM), a type of time-series statistical model in machine learning. More specifically, in this method, we use HMMs to

✉ Antoni B. Chan
abchan@cityu.edu.hk

✉ Janet H. Hsiao
jhsiao@hku.hk

¹ Department of Psychology, University of Hong Kong, Hong Kong, Hong Kong

² Department of Computer Science, City University of Hong Kong, Hong Kong, Hong Kong

³ Department of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA

⁴ The State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong, Hong Kong

directly model eye movement data, and the hidden states of the HMMs correspond to the ROIs of eye movements. The transition probabilities between hidden states (ROIs) represent the temporal pattern of eye movements between ROIs. To account for individual differences, we use one HMM to model one person's eye movement pattern in a visual task in terms of both person-specific ROIs and transitions among the ROIs (Fig. 1a). An individual's HMM is estimated from the individual's eye movement data using a variational Bayesian approach that can automatically determine the optimal number of ROIs. In addition, individual HMMs can be clustered according to their similarities (Coviello, Chan, & Lanckriet, 2014), to reveal common patterns. Differences among models can be quantitatively assessed using likelihood measures—that is, by calculating the log-likelihood of data from one model being generated by another model. Thus, this

method is particularly suitable for examining individual differences in eye movement patterns and their associations with other cognitive measures. Also, since HMM is a probabilistic time-series model, it works well with a limited amount of data (e.g., 20 trials), which is in contrast to deep learning methods that require large amounts of data to train effectively. Thus, the EMHMM approach is especially suitable for psychological research in which data availability is limited or data collection is time consuming.

We have successfully applied the EMHMM method to face recognition research and discovered novel findings thus far not revealed by existing methods. For example, we discovered two common eye movement patterns for face recognition: a “holistic” pattern in which participants mainly look at the center of the face, and an “analytic” pattern that involves more frequent eye movements between the two eyes and the face

Fig. 1 **a** The representative holistic and analytic eye movement patterns discovered through clustering 34 young and 34 older adults' eye movement patterns in face recognition using the eye movement analysis with hidden Markov model (EMHMM) approach in Chan, Chan, Lee, and Hsiao (2018). In the representative models, ellipses show regions of interest (ROIs) corresponding to 2-D Gaussian emissions. The prior values show the probabilities that a trial starts from the ellipse/ROI. The transition probabilities show the probabilities of observing a particular transition to the next ROI from the current ROI. The small images to the right of the ROI ellipses show the assignment of actual fixations to

the ROIs and the corresponding fixation heatmaps. Note that the clustering algorithm (Coviello, Chan, & Lanckriet, 2014) summarizes individual models in a cluster into a representative HMM using a prespecified number of ROIs, which may result in overlapping ROIs. Here the number of ROIs in the representative models was set to 3, the median number of ROIs in the individual models. **b** The frequencies of young and older adults adopting holistic patterns and analytic patterns: Significantly more older adults adopted holistic patterns, and more young adults adopted analytic patterns. **c** Group fixation heat maps for the older and young adults.

center (Fig. 1a). Analytic patterns were associated with better recognition performance, and this effect was consistently observed across different culture and age groups (e.g., Chan, Chan, Lee, & Hsiao, 2018; Chuk, Chan, & Hsiao, 2017; Chuk, Crookes, Hayward, Chan, & Hsiao, 2017). In addition, we found that significantly more participants (75%) used same eye movement patterns when viewing own- and other-race faces than different patterns (Chuk, Crookes, et al., 2017). In contrast, only around 60% of participants used the same eye movement patterns between face learning and recognition, and their recognition performance did not differ significantly from those using different patterns, in contrast to the scan path theory, which posits that eye movements during learning have to be recapitulated during recognition for the recognition to be successful (Chuk et al., 2017). We also found that more older adults adopted holistic patterns, whereas more young adults adopted analytic patterns (Fig. 1b). This difference was not readily observable from the group eye fixation heat maps, demonstrating the power of the EMHMM method (Fig. 1c; Chan et al., 2018). Among older adults, holistic patterns were associated with lower cognitive status as assessed using Montreal Cognitive Assessment (HK-MoCA; Yeung, Wong, Chan, Leung, & Yung, 2014), particularly in executive function and visual attention abilities (as assessed by the Tower of London and trail-making tests). Interestingly, this association was replicated when we used models of the discovered common patterns to assess new participants' eye movement patterns when viewing new face images, suggesting the possibility of developing representative models from the population for cognitive impairment screening purposes.

Although the EMHMM method has been shown to be effective in accounting for individual differences in eye movement patterns, the present method assumes that the viewer's cognitive state does not change during the task and thus the viewer's eye movement pattern throughout the task remains consistent. This assumption may not be true in tasks that involve cognitive state changes. For example, for tasks involving decision making, participants typically need to explore different options and decide which option they prefer. This process may inevitably involve at least two cognitive states: exploration and decision making. More specifically, before a decision response, participants may switch between cognitive states of exploration and decision making, during which participants may adopt different eye movement patterns. Findings from some previous studies support this speculation. For example, Shimojo, Simion, Shimojo, and Scheier (2003) conducted a two-alternative forced choice (2AFC) preference decision making task, in which participants were required to look at two face images and then to decide which one they liked more. The two faces were shown on the left and the right side of the screen. The results showed that participants spent roughly equal time exploring the two faces at the beginning, but about 800 ms before their decision

response, they spent significantly longer time looking at the image that they were about to choose. Shimojo et al. (2003) coined the term "gaze cascade effect" to describe this phenomenon. In addition, they argued that gaze shifts, particularly those at the end of a trial, were essential to not only reflecting, but also shaping preferences. In a follow-up study (Simion & Shimojo, 2007), the same setting was adopted but the face images were removed from the screen before participants indicated their choices. The results showed that participants still looked at the to-be-chosen side of the screen more often before they made their decisions even though the stimuli were absent from the screen. This finding suggests that the gaze orienting behavior in preference decision making reflects an indispensable cognitive state in the decision making process.

Indeed, previous research has reported that eye movement patterns reflect changes in cognitive states. For example, Haji-Abolhassani and Clark (2014) reported that when viewing real scene images, participants showed different eye movement patterns when being asked to make different judgments about the images, and their eye movements could be used to infer the type of judgment they were making (see also Coutrot, Hsiao, & Chan, 2018; Henderson, Shinkareva, Wang, Luke, & Olejarczyk, 2013; Kanan et al., 2015). Ahlstrom and Friedman-Berg (2006) found that levels of mental workload during an air-traffic control simulation were correlated with different eye movement patterns. Lemonnier, Brémont, and Baccino (2014) showed that in a driving simulation task, participants' decision to either stop at or drive through a junction was associated with different eye movement patterns prior to the decision. In more complicated tasks, multiple cognitive states may be involved, and thus participants' eye movement patterns may change accordingly during a trial. In the literature, visual attention or eye movement changes corresponding to transitions in cognitive states during a complicated cognitive task have been commonly modelled using probabilistic models such as HMM. For example, Liechty, Pieters, and Wedel (2003) used two hidden states to represent global and local covert attention in an HMM to model shifts between global and local covert attention when people were viewing advertisements. Simola, Salojärvi, and Kojo (2008) used an HMM to model eye movement behavior during an online information search task and discovered three hidden states that could be interpreted as scanning, reading, and decision making, respectively. Yi and Ballard (2009) used states in a dynamics Bayes network to represent subtask goals in modeling task control in sandwich making (see also Hayhoe & Ballard, 2014). Van der Lans, Pieters, and Wedel (2008) used an HMM with two hidden states that corresponded to the cognitive states of localization and identification, respectively when participants were searching for a product of a specific brand on the shelves: During localization, participants explored the environment to locate the target stimulus, whereas during identification, they examined the currently fixated

stimulus and compared it against the target information in mind. They showed that the two hidden states involved different saliency information, suggesting that during the two cognitive states, participants looked at different areas of the scenes. These findings were consistent with Shimojo et al. (2003), suggesting that multiple cognitive states may exist during a trial and that they may be associated with different eye movement patterns.

Note that in these previous studies using HMMs or probabilistic models to model eye movement or visual attention in a complicated cognitive task, the hidden states of the models represented cognitive states (e.g., Hayhoe & Ballard, 2014; Liechty et al., 2003). Thus, these models capture the temporal dynamics of cognitive state transitions but not the dynamics of eye movements. In contrast, in our present EMHMM approach, we directly use HMMs to model eye movements, and hidden states of the HMMs are directly associated with ROIs of eye movements. This allows us to discover temporal dynamics of eye movements in terms of ROIs and transitions among ROIs specific to an individual, but not temporal dynamics of cognitive state changes in a complicated cognitive task.

To better understand individual differences in eye movement patterns in real-life complicated cognitive tasks, new methods that can discover multiple cognitive states that occurred during a task and the eye movement pattern associated with each cognitive state are required. An intuitive solution to the problem is to extend the standard HMMs in the EMHMM method to hierarchical HMMs.¹ A hierarchical HMM with two layers contains a high-level HMM and several low-level HMMs. Each low-level HMM can be used to learn the eye movement pattern of a cognitive state. The high-level HMM acts like a “switch” that captures the transitions between cognitive states. It does so by learning the transitions between the low-level HMMs. We call this a *switching HMM* (SHMM). Here we aim to develop methodologies required for using an SHMM to model transitions between cognitive states and their associated eye movement patterns in a cognitive task. We then apply the new methodologies to model eye movements in a preference decision making task. More specifically, we conduct a face preference decision making task using the same paradigm as that described in Shimojo et al. (2003) and Shimojo, Simion, and Changizi (2011), in which participants were presented with two face images in each trial and asked to judge which face they liked more. On the basis of previous findings (Shimojo et al., 2003), at least two different eye movement patterns are expected to be observed during a trial: One involves no fixation preference over either stimulus, which may be related to exploration and information

sampling, whereas the other involves a higher percentage of fixations over the to-be-chosen stimulus (i.e., the gaze cascade effect), which both reflects and shapes preferences. Here we aim to use an SHMM to capture these dynamics in cognitive state and eye movement pattern. Following the existing EMHMM approach, we will train one SHMM per participant to summarize the participant’s eye movement behavior during the task. We will then cluster individual SHMMs according to their similarities to discover common eye movement patterns in the task, and examine whether different eye movement patterns are associated with different decision-making behavior. We will also compare the results using SHMM with those using standard HMM to examine the advantage of SHMM. The methodologies developed here can be applied to other tasks that involve eye movement pattern/cognitive state changes, making long-lasting impacts on how researchers across disciplines use eye movements to understand cognitive behavior. The new methodologies will be made freely available to the research community for noncommercial use under an open-source license agreement in the form of a Matlab toolbox, EMSHMM.

Method

Face preference decision-making task

Here we used eye movement data collected from a face preference decision-making task to illustrate the EMSHMM approach. We planned to conduct a comparison between two participant groups as the result of clustering using EMSHMM, and thus 24 participants were recruited according to a previous study using a similar task (Shimojo et al., 2011), in which 12 participants were recruited. In the present study, data from half of the participants were from Shimojo et al. (2011). The rest of the participants were recruited at the University of Hong Kong. Following Shimojo et al. (2011), the preference decision-making task was a 2AFC task, which contained two parts. In Part 1, participants were presented with 120 (female and male) computer generated, bald faces. They were asked to rate, from 1 to 7, how attractive the faces were. These ratings were used for pairing stimuli for Part 2 and the eye movements were not analyzed. After Part 1 was finished, the faces of the same gender that received similar ratings were paired to form 60 pairs as the stimuli in Part 2. In Part 2, in each trial, each pair of faces were shown on the screen with one on the left and one on the right. Participants were required to indicate which face they preferred. There was no time limit. Participants could move their eyes freely to compare the two images. They were told to press a button to indicate which face (left or right) they preferred once they had made their decision. Eye movements were recorded using an EyeLink 1000 eyetracker. More details about the experimental

¹ Note that there are other hierarchical HMMs for other purposes. For example, please see Camci and Chinnam (2006) and Hariri, Shirmohammadi, and Pakravan (2008) for more details.

task can be found in Shimojo et al. (2003) or Shimojo et al. (2011). In data acquisition, we extracted fixation location information using EyeLink Data Viewer. Saccade motion threshold was 0.15 deg of visual angle; saccade acceleration threshold was 8,000 deg/s²; saccade velocity threshold was 30 deg/s. These are the EyeLink defaults for cognitive research.

Switching hidden Markov model

A standard hidden Markov model (HMM) contains a vector of prior values of the hidden states, a transition matrix of the hidden states, and a Gaussian emission for each hidden state. The prior values indicate the probabilities of the time-series data starting from the corresponding hidden states. The transition matrix indicates the transition probabilities between any two hidden states. The Gaussian emissions indicate the probabilistic associations between the observed time-series data and the hidden states. In our EMHMM approach (Chuk et al., 2014), the hidden states correspond to the ROIs, the emissions are the eye fixation locations, and the emissions in an ROI are represented as a 2-D Gaussian distribution (Fig. 1a).

In contrast to a standard HMM, an SHMM contains two levels of hidden state sequences; the low-level hidden state sequence models the temporal pattern of the time-series data following a standard HMM, and the high-level hidden state sequences indicate the transitions between the HMM parameters used by the low-level hidden state sequence. Formally, let $z_{n,t} \in \{1, \dots, K\}$ be the low-level hidden states, and $s_{n,t} \in \{1, \dots, S\}$ be the high-level hidden states, and $x_{n,t}$ be the observations, where n indexes the sequences and t indexes time. Both the high-level state sequence and the low-level

state sequence are first-order Markov chains. The prior probability and transitions of the low-level hidden state depends on the current high-level state,

$$p(s_n) = p(s_{n,1}) \prod_{t=2}^{\tau_n} p(s_{n,t} | s_{n,t-1}) \quad (1)$$

$$p(z_n | s_n) = p(z_{n,1} | s_{n,1}) \prod_{t=2}^{\tau_n} p(z_{n,t} | z_{n,t-1}, s_{n,t}), \quad (2)$$

where τ_n is the length of the n th sequence. The high-level state sequence is parameterized by the prior vector ρ and transition matrix B ,

$$p(s_{n,1} = j) = \rho_j, p(s_{n,t} = j' | s_{n,t-1} = j) = b_{j,j'}. \quad (3)$$

Given that the high-level state is $s_{n,t} = j$, the low-level state sequence is parametrized by the prior vector $\pi^{(j)}$ and transition matrix $A^{(j)}$,

$$p(z_{n,1} = k | s_{n,1} = j) = \pi_k^{(j)}, p(z_{n,t} = k' | z_{n,t-1} = k, s_{n,t} = j) = a_{k,k'}^{(j)}. \quad (4)$$

The emission densities are Gaussians and depend only on the low-level hidden state (i.e., they are shared among high-level states),

$$p(x_{n,t} | z_{n,t} = k) = N(x_{n,t} | \mu_k, \Lambda_k^{-1}), \quad (5)$$

where μ_k, Λ_k^{-1} are the mean vector and covariance matrix of the Gaussian. Here we assume that the number of low-level states for each high-level state is the same. The joint probability model for the SHMM is

$$p(\mathbf{X}, \mathbf{Z}, \mathbf{S}) = \prod_{n=1}^D [p(s_{n,1}) p(z_{n,1} | s_{n,1}) p(x_{n,1} | z_{n,1}) \prod_{t=2}^{\tau_n} p(s_{n,t} | s_{n,t-1}) p(z_{n,t} | z_{n,t-1}, s_{n,t}) p(x_{n,t} | z_{n,t})]. \quad (6)$$

In practice, the SHMM can be turned into a standard HMM by combining the high-level and low-level hidden state variables into a single hidden-state variable, whose values are the Cartesian product of the low- and high-level state values. Here we assume the low-level states are shared among the high-level states (S), and thus the number of low-level states (K) is the same for each high-level state. Consequently, the equivalent standard HMM has $S \times K$ augmented hidden states $\tilde{z}_{n,t}$. The augmented states take a value pair (j, k) , where j indicates the high-level state and k indicates the low-level state. The transition probabilities and the prior values therefore are defined as

$$p(\tilde{z}_{n,t} = (j', k') | \tilde{z}_{n,t-1} = (j, k)) = \tilde{a}_{(j,k),(j',k')} = b_{j,j'} a_{k,k'}^{(j)}, \quad (7)$$

$$p(\tilde{z}_{n,1} = (j, k)) = \tilde{\pi}_{(j,k)} = \rho_j \pi_k^{(j)}. \quad (8)$$

Thus, the relationship between the augmented hidden states and the two separate levels of hidden state sequences are defined as

$$\begin{aligned} p(\tilde{z}_n) &= p(z_n, s_n) = p(z_n | s_n) p(s_n) \\ &= p(z_{n,1} | s_{n,1}) p(s_{n,1}) \prod_{t=2}^{\tau_n} p(z_{n,t} | z_{n,t-1}, s_{n,t}) p(s_{n,t} | s_{n,t-1}) \\ &= p(s_{n,1}, z_{n,1}) \prod_{t=2}^{\tau_n} p(s_{n,t}, z_{n,t} | s_{n,t-1}, z_{n,t-1}) \end{aligned} \quad (9)$$

The transition matrix and the prior vector have block structures,

$$\tilde{\mathbf{A}} = \begin{bmatrix} b_{1,1}\mathbf{A}^{(1)} & b_{1,2}\mathbf{A}^{(2)} \\ b_{2,1}\mathbf{A}^{(1)} & b_{2,2}\mathbf{A}^{(2)} \end{bmatrix}, \quad (10)$$

$$\tilde{\boldsymbol{\pi}} = \begin{bmatrix} \rho_1 \boldsymbol{\pi}^{(1)} \\ \rho_2 \boldsymbol{\pi}^{(2)} \end{bmatrix}. \quad (11)$$

In our implementation, the high-level hidden states represent the cognitive states, whereas the low-level hidden states correspond to ROIs over the stimuli. The high-level state sequence has its own transition matrix, which governs the transitions between cognitive states. The low-level states (ROIs) are shared among the high-level states. Taking the preference decision-making task as an example, here we assumed that participants have two cognitive states: an exploration period that involves information sampling without preference to a specific stimulus, and a preference-biased period during which preference is formed and eye movement behavior can be influenced by the preference. We assumed a simplified decision process in which the participant starts in the exploration period, and that once the participant has sampled enough information in the exploration period, they transition to the preference-biased period and cannot transition back to the exploration period. That is, once the preference-biased period is entered it cannot be exited until the response decision. The gaze cascade effect observed in Shimojo et al. (2003) showed an increased percentage of eye fixations on the stimulus to be chosen prior to the response. To focus on modeling this effect, we assumed that the low-level HMM had two ROIs, each corresponded to a stimulus for choice. Figure 2 illustrates an example model summarizing a participant's eye movement pattern. As is shown in the figure, the high-level HMM consisted of two cognitive states as its hidden states: exploration and preference-biased periods. The blue arrows indicate the transitions between the two states, and the numbers indicate transition probabilities. The eye movements within each

state were modeled with a low-level HMM, whose hidden states represent the ROIs of eye movements. The red arrows represent transitions between the ROIs. The two cognitive states have the same ROIs but different transition probabilities.

Training individual SHMMs

We performed the expectation-maximization (EM) algorithm to estimate the SHMM parameters (see the Appendix for details). In the E-step, the responsibilities were calculated using the standard forward-backward algorithm with the block transition matrix, initial state vector, and emission densities. In the M-step, the prior and pairwise responsibilities were summed over the high-level and the low-level states, respectively to yield the parameter updates for both the high-level states and the low-level states.

For instance, the prior responsibilities were summed over the low-level hidden states for each of the high-level state to yield the parameter updates for the low-level state sequence, and then they were summed over the high-level states to yield the parameter updates for the high-level state sequence. Similarly, the pairwise responsibilities were summed over the low-level hidden states for each high-level state to yield the updates for each low-level transition, and then were summed over the high-level hidden states to yield the updates for the switching (transition) matrix of the high-level state sequence.

For each participant, we trained two SHMMs, one using the data from the left-selected trials and one using the data from the right-selected trials. We combined the two SHMMs into one for each individual as follows. Preliminary analysis indicated that the exploration periods of the left-selected and right-selected trials were similar. In other words, the eye movements during exploration periods were consistent regardless of which side was selected. Thus, the transition matrices for the exploration period of the left-selected and right-

Fig. 2 Illustration of an example switching hidden Markov model (SHMM) summarizing a participant's eye movement pattern in the preference decision-making task in Shimojo et al. (2003). The blue arrows

indicate the transition probabilities between the cognitive states (high-level states); the red arrows indicate the transition probabilities between the regions of interest (ROIs; low-level states).

selected models were directly averaged together. For the preference-biased period, we averaged the two preferred-side parameters and the two non-preferred-side parameters, essentially normalizing the right-selected trials into left-selected trials. To focus the analyses on the transition between the stimuli during preference decision making, we used only two Gaussian emissions per model, one on each side, for the low-level states (Fig. 2). For SHMM estimation, we initialized one Gaussian centered over each stimulus, with covariance that covered the stimulus. Thus, it could be considered here that the low-level states of the model were prespecified (and thus not hidden), since with good confidence we could determine which stimulus was viewed. The advantage of using Gaussian rather than discrete emissions is that it can be easily extended to analyses that explore more ROIs (i.e., more hidden lower-level states) on each stimulus. We used two high-level hidden states to reflect that the participants would transition from the exploration period to the preference-biased period during a trial.

For SHMM estimation, we initialized the transition matrices of the high-level state as $[.95, .05; 0.0, 1.0]$ and high-level prior as $[1.0, 0.0]$, which encodes the assumed behavior of starting in the exploration period and staying there (.95), and then eventually transitioning to the preference-biased period (.05) and not back (0.0). During training, this initialization causes the probability of transitioning from the preference-biased to the exploration period to stay at 0 (since all potential sequences that transition from the preference-biased to the exploration period are given probability zero). The transition matrices of the low-level states were initialized as uniform distributions. After the initialization, the Gaussian ROIs and transition probabilities were updated in the EM algorithm.

Clustering to discover common patterns

To examine the general eye movement pattern adopted by all individuals during the exploration period, we created an HMM using the exploration transition matrix and Gaussian emissions for each individual, and clustered (or summarized) these HMMs into one group using the variational hierarchical expectation maximization (VHEM) algorithm (Coviello et al., 2014). The VHEM algorithm clusters HMMs into a predefined number of groups according to their probability distributions and characterizes each group using a representative HMM. A similar procedure was performed to obtain the general eye movement pattern for the preference-biased period.

To examine individual differences in decision-making behavior, we clustered the participants' high-level transition matrices using the *k*-means clustering algorithm (MacQueen, 1967) into two groups. Here we aimed to discover differences in the high-level cognitive behavior involving the explorative and preference-biased periods. For each group, we then

computed representative exploration-period and preference-biased-period HMMs by running the VHEM algorithm (Coviello et al., 2014) on the exploration-period and preference-biased period HMMs of that group, respectively.² We then examined how participants in the two groups differed in decision making behavior, including their differences in the gaze cascade plot/effect (e.g., Shimojo et al., 2003), transition between the exploration and preference-biased periods, distribution of number of fixations in exploration/preference-biased periods, and accuracy of inferring preference choices from eye movement data.

Transition between exploration and preference-biased periods

After we clustered participants' high-level transition matrices into two groups, we further investigated how the participants in these two groups differ in decision-making behavior. More specifically, we intended to reveal, from the beginning to the end of a trial, the probability that the participant was in the preference-biased period. For each trial, we computed the posterior probabilities of all possible high-level state sequences given the observed eye gaze data. We then computed a sequence of probabilities of being in the preference-biased period by computing the expectation over the high-level state sequences—that is, by computing the weighted average over all high-level state sequences in which the weights are the posterior probabilities. This was performed on all trials in all participants. Since the duration differed in each trial, to examine the proportion of time a participant spent in the exploration and preference-biased periods relative to the whole trial, we normalized the different durations across the trials by dividing each trial into the same number of segments (21 in our experiments). Then, we calculated, for each segment in a trial, the probability of time that the participant was in the preference-biased period. For each participant, we calculated the mean probability across all trials for each segment. Finally, we averaged the means across participants in the same group and plotted the mean probability at each trial segment for the two participant groups separately. The plot represented the percentage of time at each trial segment the participants were in the preference-biased period. This examination allowed us to see the difference in temporal dynamics of cognitive state changes in a trial between groups.

We also examined the number of fixations in the exploration and preference-biased period between the two groups. To this end, for each trial and each participant, we computed the posterior probabilities of all high-level state sequences, and then counted the numbers of fixations in the exploration and preference-biased periods. The aggregated probabilities over

² Since we were interested in the differences in the transitions between the two sides, we forced all HMMs to use the same set of ROIs that covered each face.

all trials and all high-level sequences was then used to form a probability distribution of number of fixations in each high-level state for a participant. The participants' probability distributions were then averaged together in each group.

Inference of individual preferences

We also examined whether SHMMs can be used to infer an individual's preference choice in a trial. For each participant, we split the trials in the preference decision making task into two sets: one for the trials in which the left-side image was chosen to be preferred, and the other for the trials in which the right-side image was chosen. Since in the preference decision-making task, face images used in a trial were matched in gender and attractiveness ratings. Thus, the two sides were expected to be chosen equally often. For each set, we used all but one trial to train a left-selected and a right-selected SHMM and used the held-out trial for testing. For testing, we created an aggregated SHMM from the two trained SHMMs, which could be used to infer the participant's choice. In particular, the aggregated SHMM contained three high-level states: exploration, left-selected preference-biased, and right-selected preference-biased periods. In the high-level transition matrix of this aggregated SHMM, the transition probability of moving from the exploration period to a preference-biased period was equally divided between the left-selected and right-selected preference-biased periods. Finally, to perform inference of the participant's choice on the test eye fixation sequence, the posterior probability of the high-level state of the last fixation $p(s_T | x_1, \dots, x_T)$ given the test sequence $(x_1, \dots, x_T)$ was computed, which indicates the probability of being in either the left-selected or right-selected preference-biased period at the end of the trial. The left/right-selected preference-biased period with higher probability was predicted as the choice. This was repeated over all trials for each participant to calculate the inference accuracy.

We performed the inferences in three ways. First, to examine the percentage of fixations in a trial required for making above-chance-level inferences, we began with using the first 10% of the fixations for the inferences and increased this proportion by 5% each time until all fixations (100%) were used. The inference task was therefore conducted 19 times, and we calculated the mean inference accuracy each time. Second, to examine how quickly a decision could be inferred, we performed inference on increasing duration of fixation sequences from the beginning of the trial (e.g., the fixations in the first 1 s, the fixations in the first 2 s, etc.). Third, the gaze cascade model suggested that although preferences were shaped during a trial when participants switched their fixations between the two stimuli, the fixations immediately before the end of the trial were usually significantly biased to the preferred stimulus. Thus, these fixations should be better predictors for participants' preference than the earlier fixations.

Accordingly, in a separate examination we used only the fixations in the last 2 s before the decision to perform the inferences.

Results

Categorization of individual SHMMs

We trained one SHMM for each participant and created (1) a standard HMM using the exploration period transition matrix to represent the participant's eye movement pattern during the exploration period and (2) a standard HMM using the preference-biased period transition matrix to represent the participant's eye movement pattern during the preference-biased period.

First, Table 1 shows the average high-level state transition matrix. Participants started in the exploration period and had a 55% probability to remain in that period. Otherwise, there was a 45% probability to transition to the preference-biased period, where participants remained until the end of the trial.

Next, we clustered the 24 participants' exploration period HMMs into one representative HMM using the VHEM algorithm. Table 2 shows the transition matrix of the representative exploration period HMM. The output showed that participants tended first to view the left side with several fixations, and then transition to viewing the right side for several fixations. After viewing the right side, the participants rarely looked back at the left side (12% probability). This suggests that participants performed a quick scan of the two sides during the exploration period.

Similarly, we clustered the 24 participants' preference-biased period HMMs into one representative HMM using the VHEM algorithm. Table 3 shows the transition matrix of the representative preference-biased period HMM.

It can be seen that participants in the preference-biased period were biased to remain looking at the to-be-chosen side (77%) more often than to the not-chosen side (67%). Furthermore, they were more likely to transition from the not-chosen side to the chosen side (33%) than vice versa (23%). This was consistent with the gaze cascade effect found in previous studies (e.g., Shimojo et al., 2003).

Next, to investigate individual differences in the gaze cascade effect, we clustered participants on the basis of their high-level transition matrices into two groups. We found that

Table 1 High-level state transition matrix of all participants

	Exploration	Preference
Prior	1.00	.00
Exploration	.55	.45
Preference	.00	1.00

Table 2 Transition matrix of the representative exploration HMM obtained by clustering the 24 exploration period HMMs into one group

	Left	Right
Prior	.70	.30
Left	.64	.36
Right	.12	.88

one group (Group A) included 11 participants, and the other (Group B) included 13 participants. Table 4, 5 and 6 show the high-level transition matrices and the transition matrices of the representative exploration and preference-biased period HMMs for the two groups.

From the high-level state transition matrix, Group A had a higher probability to stay in the exploration period (68%) than did Group B (45%). Thus, the exploration period was longer for Group A than for Group B. By examining the transition matrix during the exploration period, we found that Group A had a stronger tendency to start exploring from the left side (76%) than did Group B (64%) and had a higher probability to remain looking at the left side (67%) than did Group B (60%). After switching to the right side, Group A also had a higher probability to transition back to the left side (17% vs. 9%).

During the preference-biased period, the participants in Group A showed an apparent fixation bias to remain looking at the chosen side. More specifically, the participants in Group A had a stronger tendency to keep looking at the to-be-chosen side (83%) than did Group B (71%), and switched less often between the two sides than did Group B.

Cascade plot

Our analyses showed that participants were biased to look more at the side that they were about to choose during the preference-biased period, which was consistent with the gaze cascade effect observed in the literature. However, our clustering results showed that this difference was more obvious for one group of participants (Group A) than for the other (Group B). To visualize the difference in the gaze cascade effects between the two participant groups, we generated a gaze cascade

Table 3 Transition matrix of the representative preference-biased HMM obtained by clustering the 24 preference-biased period HMMs into one group

	To Chosen	To Not-Chosen
Prior	.53	.47
From Chosen	.77	.23
From Not-Chosen	.33	.67

Table 4 Transitions matrices of the high-level states of Group A (11 participants) and Group B (13 participants)

	Exploration	Preference
Group A		
Prior	1.00	.00
Exploration	.68	.32
Preference	.00	1.00
Group B		
Prior	1.00	.00
Exploration	.45	.55
Preference	.00	1.00

plot like that seen in Shimojo et al. (2003). The plot showed the probability that participants were looking at the image to be chosen during the last 2.5 s prior to the response. Figure 3 shows the gaze cascade plots of the two groups and their average.

It can be seen from the “All” plot that participants spent more time on inspecting the side that they were about to choose near the end of a trial. The proportion of time spent on the chosen side went from chance level .5 steadily up until it reached around .87. We estimated the probability that each participant looked at the chosen stimuli at each time point at a 100-ms interval. To test the hypothesis that across the time intervals there was significant difference in the probabilities of looking at the chosen side, and that the two groups of participants differed in their time interval effects, we performed a mixed analysis of variance (ANOVA) on the probabilities of looking at the chosen side, with time interval as a within-subjects variable and group as a between-subjects variable. The results showed a significant main effect of time interval, $F(3.043, 66.938) = 52.163$, $p < .001$, $\eta_p^2 = .703$; a significant main effect of group, $F(1, 22) = 5.481$, $p = .029$, $\eta_p^2 = .199$; and a marginal interaction between time interval and group, $F(3.043, 66.938) = 2.307$, p

Table 5 Transition matrices of the representative exploration periods of Group A (11 participants) and Group B (13 participants)

	Left	Right
Group A		
Prior	.76	.24
Left	.67	.33
Right	.17	.83
Group B		
Prior	.64	.36
Left	.60	.40
Right	.09	.91

Table 6 Transition matrices of the representative preference-biased periods of Group A (11 participants) and Group B (13 participants)

	Chosen	Not Chosen
Group A		
Prior	.50	.50
Chosen	.83	.17
Not Chosen	.25	.75
Group B		
Prior	.54	.46
Chosen	.71	.29
Not Chosen	.39	.61

$= .084$, $\eta_p^2 = .095$.³ In addition, we observed a significant linear trend, $F(1, 22) = 126.657$, $p < .001$, $\eta_p^2 = .852$, and a quadratic trend, $F(1, 22) = 19.609$, $p < .001$, $\eta_p^2 = .471$, across time intervals. These results showed that during the last 2.5 s before response, participants had a significant increase in probability of looking at the chosen side—that is, the cascade effect. In addition, Group A had higher a probability of looking at the chosen side than Group B, indicating a stronger cascade effect.

Post-hoc t tests showed that participants started to look at the chosen item significantly above chance level around 1,100 ms before the end of the trial (i.e., the beginning of the gaze cascade effect), $t(23) = 2.27$, $p = .033$, $d = 0.46$ (here, d refers to Cohen's d , an effect size measure to indicate a standardized difference between two means), until the end of the trial. Within this time period, the probability to look at the chosen face rose from 58% to 87%. Furthermore, there was a short period, from 2,200 to 1,600 ms before the end of the trial, in which the participants also looked at the chosen side with slightly higher probability than chance level, according to a t test (55%; Fig. 3).

The plots of the two participant groups show some interesting differences. The participants in Group A showed a stronger gaze cascade effect. The probability that Group A looked at the chosen item reached 94.5% at the end. A t test was conducted to examine when their probability of viewing the chosen stimulus was above chance level. The result showed that this occurred around 1,000 ms before the end of a trial, $t(10) = 2.44$, $p = .035$, $d = 0.74$, at which time point they spent about 66% of their time on the chosen item. In contrast, the participants in Group B had a weaker cascade effect, looking at the chosen item about 81% of the time at the end. A t test indicated that the proportion of time spent on the chosen item was significantly above the chance level 900 ms before the end of a trial, $t(12) = 2.29$, $p = .041$, $d = 0.64$, from

which point the probability to view the chosen side increased from 59% to 81%. Group B also exhibited a short period of slightly higher than chance viewing (54%) of the chosen side between 2,100 and 1,600 ms before the end of the trial (Fig. 3). We also compared the proportions at each time point between the two groups using independent-sample t tests. We found that from 700 ms before the end of a trial to the end of the trial, Group A spent significantly more time looking at the to-be-chosen item than did Group B (Fig. 3). Thus, although both groups exhibited the gaze cascade effect, their effects differed in both magnitude and onset time, suggesting substantial individual differences in the gaze cascade effect. Our ESMHMM method allows us to reveal these individual differences in the gaze cascade effect through clustering participants' eye movement patterns according to their similarities.

In addition, with the HMM-based approach, we can quantitatively assess the similarity of each participant's eye movement pattern during the preference-biased period to the representative pattern of Group A or Group B by calculating the log likelihood of the participant's eye movement pattern being generated by the representative model. To quantify a participant's eye movement pattern along the continuum between the representative patterns of Groups A and B, we defined the A–B scale as $(LA - LB)/(|LA| + |LB|)$, where LA is the log likelihood of the eye movement pattern being generated by the Group A model, and LB is the log likelihood of the eye movement pattern being generated by the Group B model (Chan et al., 2018). The larger the A–B scale, the more similar the pattern is to the representative pattern of Group A, and the smaller the A–B scale, the more similar the pattern is to the representative pattern of Group B. Among our participants, we observed a significant positive correlation between A–B scale and the gaze cascade effect as measured by the average probability of looking at the chosen item from the onset of the effect (1,000 ms prior to the response) to the end ($r = .50$, $p = .012$), or in other words, the more similar a participant's eye movement pattern during the preference-biased period was to the representative pattern of Group A, the stronger the gaze cascade effect.

Transition between exploration and preference-biased periods

To examine whether the two participant groups differed in their transition behavior between the two cognitive states throughout a trial, we normalized trial duration by dividing each trial into 21 time segments and examined, for each participant, the percentage of trials on which, or the frequency with which, the participant was in the preference-biased period during each time segment (see the Method section for details). Figure 4a shows the average probabilities of the two groups being in the preference-biased period throughout a trial.

³ Greenhouse–Geisser correction was applied whenever the assumption of sphericity was not met.

Fig. 3 Gaze cascade plots during the last 2.5 s before response, for the two participant groups (Groups A and B) and all participants (All). The red, blue, and green stars on the top indicate the time points at which the gaze cascade effect occurred (i.e., the probability of looking at the chosen

item was significantly higher than the chance level, based on t test) for each group and for all participants. The black stars at the bottom indicate the time points during which the two groups had significant differences in their proportions of time spent on the chosen items.

The results showed that for both groups, the probability of being in the preference-biased period increased soon after a trial began. To test the hypothesis that Groups A and B had different probabilities across the time segments, a mixed ANOVA on probabilities of being in the preference-biased period, with time segment as a within-subjects variable and group as a between-subjects variable, was conducted. The results showed a significant time segment effect, $F(1.384, 30.445) = 339.091$, $p < .001$, $\eta_p^2 = .939$; a significant group effect, $F(1, 22) = 7.770$, $p = .011$; and a significant interaction between group and time segment, $F(1.384, 30.445) = 10.064$, $p = .001$, $\eta_p^2 = .314$. These results indicated that, in general, Group B had a higher probability of being in the preference-biased period than did Group A, especially in the beginning time segments of a trial (Fig. 4a). Note that although Group A entered the preference-biased period later, they had a stronger gaze cascade effect. In contrast, Group B entered the preference-biased period earlier but had a weaker gaze cascade effect.

A t test was used to test the hypothesis that Group A and Group B differed in their average numbers of fixations per trial. We found that Group A ($M = 29.0$) had a significantly larger average number of fixations in a trial than did Group B ($M = 13.3$), $t(22) = 4.44$, $p < .001$, $d = 1.75$. Group A also had larger average numbers of fixations in the exploration period (Group A, $M = 3.38$; Group B, $M = 1.88$), $t(22) = 8.46$, $p < .001$, $d = 3.37$, and in the preference-biased state (Group A, $M = 25.61$; Group B, $M = 11.5$), $t(22) = 4.10$, $p < .001$, $d = 1.62$ (Fig. 4b). Note, however, that after normalizing for the total number of fixations, Group A ($M = .147$) and Group B ($M = .173$) did not differ in the fractions of their fixations in the exploration or preference-biased period, $t(22) = -1.11$, $p = .28$, $d = 0.45$.

We also tested the hypothesis that Group A and Group B differed in their numbers of eye gaze switches between stimuli during a trial. Group A ($M = 5.66$) had more switches than

Group B ($M = 4.10$) on average, $t(22) = 2.81$, $p = .01$, $d = 1.12$. However, this effect was due mainly to Group A having more fixations in a trial in general. After normalizing for the total number of fixations in a trial, Group A ($M = .211$) had a smaller fraction of eye fixations that involved switching between stimuli than did Group B ($M = .328$), $t(22) = -7.20$, $p < .001$, $d = 2.97$. This effect suggested that Group A tended to explore a stimulus longer before switching to the other stimulus than did Group B. Note that since the clustering was based only on participants' cognitive (high-level) state transitions, these differences in number of fixations per trial and frequency of switches between stimuli emerged naturally as a result of the clustering.

Inference of participants' preference choices

We then explored whether the models could be used to infer an individual's preference choice in each trial given partial eye movement data (please refer to the Method section for details). We started by using the fixations from the first 10% of the trial according to the normalized duration, and gradually increased this percentage in steps of 5%. Figure 5 shows the average inference accuracies of the two groups.

As is shown in Fig. 5a, when using the first 75% to 90% of the fixations in a trial, the average inference accuracy was higher for Group B than for Group A. In contrast, when the first 95% or all of the fixations of a trial were used, the inference accuracy for Group A was higher than that for Group B. To test the hypothesis that Group B had higher inference accuracy when 75% to 90% of the fixations were used, whereas Group A had higher inference accuracy when 95% to 100% of fixations were used, we conducted a mixed ANOVA on inference accuracy, with amount of fixations (75% to 100%, in steps of 5%) as a within-subjects variable and group as a between-subjects variable. The results showed a main effect

Fig. 4 **a** Probabilities of the participants in the two groups being in the preference-biased period at each normalized trial segment. **b** Probability distributions of number of fixations in the exploration period and number

of fixations in the preference-biased period. The vertical bars indicate standard errors.

of amount of fixations, $F(1.356, 29.840) = 30.778$, $p < .001$, $\eta_p^2 = .583$, and an interaction between amount of fixations and group, $F(1.356, 29.840) = 5.657$, $p = .016$, $\eta_p^2 = .205$. The main effect of group was not significant, $F(1, 22) = 0.077$, n.s. These results indicated that, as is shown in Fig. 5a, Group B had higher inference accuracy when a smaller amount of fixations were used (75% to 90%), whereas Group A had a higher inference accuracy when a larger amount of fixations were used. When we compared the inference accuracies against the chance level (.5) using a t test, we found that when using the first 75% to 85% of the fixations, Group B's inference accuracy was significantly above the chance level, whereas Group A's inference accuracy was not. When using the first 90% to 100% of the fixations, inference accuracy was significantly above the chance level for both groups (Fig. 5a). In other words, the participants in Group B revealed their biases to the preferred stimulus earlier than those in Group A.

To examine the actual time when a group's accuracy became significantly above chance, Fig. 5b plots the average inference accuracy using time windows (in seconds) starting

from the beginning of the trial. After the start of the trial, Group B's inference accuracy increased more rapidly than Group A's, became significantly above chance at around 3 s, and then saturated at around 4 s. In contrast, Group A's inference accuracy increased slowly, became above chance level at around 6 s, and saturated at around 10 s. Although Group A's accuracy increased more slowly, it saturated at a higher value than Group B's.

Last 2 s before response

The gaze cascade effect (Shimojo et al., 2003) suggested that eye movement patterns immediately before a preference decision response may provide a strong cue for inferring the preference. According to Fig. 3, participants started to show a tendency to look at the chosen side more often around 2 s before response. Thus, here we examined the accuracy of inferring participants' preferences using the fixations during the last 2 s before the response.

Fig. 5 **a** Average inference accuracies of the two groups using partial fixations in a trial, selected as a percentage of each trial's duration from the beginning. **b** Average inference accuracies of the two groups using different window lengths starting from the beginning of the trial (note that participants had different trial lengths, with the average length being 5.65

± 2.57 s; Group A had an average trial length of 7.67 ± 2.32 s, whereas Group B had an average trial length of 3.93 ± 1.13 s). The red, blue, and green stars at the top of each panel indicate the data points at which the accuracy was significantly higher than chance level based on *t* test, for each group and for all participants, respectively.

We first tested the hypothesis that for the participants in both Group A and Group B, the average accuracy of inferring their preference decisions using the final 2 s was significantly above the chance level, and the results supported this hypothesis (Fig. 6a): Group A, $M = .93$, $t(10) = 15.89$, $p < .001$, $d = 4.79$; Group B, $M = .71$, $t(12) = 3.79$, $p = .003$, $d = 1.05$. In addition, we tested the hypothesis that the inference accuracy of Group A was significantly higher than that of Group B. The results also supported this hypothesis, $t(22) = 3.26$, $p = .004$, $d = 1.33$. In addition, the more similar that participants' eye movement patterns during the preference-biased period were to the representative pattern of Group A, as opposed to that of Group B (as measured by the A–B scale), the higher the inference accuracy ($r = .47$, $p = .02$; Fig. 6b). This result was consistent with the observation that Group A exhibited a stronger gaze cascade effect (Fig. 3). Note that the clustering of the two groups was based completely on the eye movement data, and thus the group difference in inference accuracy emerged naturally as a result of the clustering.

Would a regular HMM, without inferring participants' cognitive state transitions, such as that used in our EMHMM approach (Chuk et al., 2014), also be able to reveal participants' preference choices? To test this, we performed the same inference task using regular HMMs in the EMHMM

approach. As is shown in Fig. 7, the average inference accuracy was higher for SHMMs ($M = .81$) than for HMMs ($M = .64$) for all participants, $t(46) = 2.71$, $p = .009$, $d = 0.78$. We conducted a two-way ANOVA on inference accuracy, with group and model (SHMM/HMM) as independent variables. The results showed a significant main effect of group, $F(1, 44) = 5.04$, $p = .03$, $\eta_p^2 = .096$, and a significant main effect of model, $F(1, 44) = 8.75$, $p = .005$, $\eta_p^2 = .166$. There was no interaction between group and model, $F(1, 44) = 1.91$, $p = .17$, however. This result demonstrated again the advantage of EMSHMM for modeling eye movement patterns in tasks that involve cognitive state changes.

Discussion

Here we have proposed a novel method, EMSHMM, for modeling eye movement patterns in tasks that involve cognitive state changes. Similar to our previous hidden Markov modeling approach to eye movement data analysis, EMHMM, the EMSHMM approach has several advantages over traditional eye movement data analysis methods, such as ROI or fixation heat map analysis, including the ability to account for individual differences in both the spatial and

Fig. 6 **a** Average inference accuracies of the two groups using fixations in the last 2 s until the response. **b** The more similar participants' eye movement patterns were to the representative pattern of Group A (i.e., the farther to the right on the x -axis), the higher the inference accuracy.

temporal dimensions of eye movements (i.e., through discovering personalized ROIs and transition probabilities among the ROIs) and to quantitatively assess these differences. In contrast to EMHMM, which uses a single regular HMM to

model eye movements and assumes a participant's strategy is consistent throughout a trial, the EMSHMM approach uses multiple low-level HMMs corresponding to different strategies/cognitive states, and a higher-level state sequence

Fig. 7 Average inference accuracy for all participants using switching hidden Markov models (SHMMs) and regular HMMs when using the last 2 s of the trials.

to capture the transitions among different strategies/cognitive states. Thus, it is especially suitable for analyzing eye movement data in complex tasks that involve cognitive state changes such as decision-making tasks.

To demonstrate the advantages of using the EMSHMM approach, we conducted a preference decision-making task, in which participants viewed two faces with similar attractiveness ratings and decided which one they preferred. Previous studies (e.g., Shimojo et al., 2003; Simion & Shimojo, 2007) reported that participants showed different eye movement patterns at different states of a trial: They usually began with exploring both alternatives, and then focused on the one they preferred by the end of the trial. These two eye movement patterns may be associated with different cognitive states. In our EMSHMM approach, we assumed that the two eye movement patterns were associated with two cognitive states, exploration period and preference-biased period, respectively. We used a switching HMM (SHMM) to summarize a participant's eye movement pattern in the preference decision-making task. The SHMM contained two ROIs that corresponded to the two faces of choice; two low-level HMMs that summarized the eye movement patterns during the exploration and preference-biased periods, respectively; and a high-level state sequence that captured the transitions between the two cognitive states.

A summary of all participants' high-level/cognitive state transitions showed that, on average, they had a 55% probability to remain in the exploration period and a 45% probability to transition to the preference-biased period, where they remained until the end of the trial (Table 1). When we summarized all participants' exploration period HMMs into one representative model, we found that participants had a bias to start from looking at the stimulus on the left side and to remain exploring there, and only then to switch to the right side (Table 2). In contrast, when we summarized all participants' preference-biased period HMMs into one representative model, we found that participants looked more often at the to-be-chosen, preferred stimulus (Table 3). These findings were consistent with the gaze cascade effect reported in previous studies (e.g., Shimojo et al., 2003; Simion & Shimojo, 2007). Indeed, when we plotted the percentage of time during which participants looked at the to-be-chosen stimulus before the end of a trial, it showed a steady increase from about 1.5 s before trial's end, demonstrating a gaze cascade effect (Fig. 3).

When we clustered participants' SHMMs into two groups according to their cognitive state transitions, we found that one group (Group A) showed a stronger and earlier gaze cascade effect than the other group (Group B; Fig. 3). The two groups also showed interesting differences in the temporal dynamics of eye movement patterns throughout a trial. More specifically, the participants in Group A entered the preference-biased period later than did Group B (Fig. 4), but they had a stronger cascade effect. In addition, Group A's

preference over the two alternatives could not be inferred with above-chance performance using early fixations of a trial until the first 90% of fixations had been used. In contrast, Group B's preference could be inferred with above-chance accuracy from only the first 75% of fixations (Fig. 5a). However, when we used only the fixations during the final 2 s before the decision response, we inferred Group A's preference with a higher accuracy than Group B's. This phenomenon showed that although the participants in Group A revealed their preference in their eye movement patterns later in a trial than did those in Group B, their eye movement patterns contained more information for inferring their preferences. Recent research has suggested that indecisiveness, or decisional procrastination, is associated with informational tunnel vision (e.g., Ferrari & Dovidio, 2000; Rassin, Muris, Booster, & Kolsloot, 2008): Indecisive individuals tend to gather more information about the item that is eventually chosen, while ignoring information about other alternatives. Accordingly, it is possible that the participants in Group B exhibited a more "indecisive" eye movement pattern than those in Group A, since they entered the preference-biased period earlier, spent proportionally less time in the exploration period, and switched more frequently between the two stimuli for choice, which may be characteristics of informational tunnel vision. Future work will examine this possibility through investigating the relationship between eye movement pattern similarity (as assessed using the EMHMM/EMSHMM approach) to the representative pattern of Group B and the participants' personality measures related to indecisiveness (e.g., Germeijs & De Boeck, 2002).

These individual differences in eye movement pattern and cognitive style during decision making had not been reported before in the literature. More specifically, previous studies only observed that decisions were related to the final fixations in a trial as revealed in the gaze cascade effect. Our analysis results showed that participants' preference could be inferred significantly earlier than the exhibition of the gaze cascade effect, and for some participants (i.e., Group B) this inference could achieve above-chance-level performance with only the first 75% of the fixations. Interestingly, these participants also tended to show a weaker gaze cascade effect. These findings demonstrated the importance of taking individual differences into account when trying to understand human decision-making behavior. Note also that although using regular HMMs without cognitive state transitions in our previous EMHMM approach could also account for individual differences in eye movement patterns, the accuracy in inferring participants' preference choices using EMSHMM was significantly higher than that using EMHMM. This phenomenon suggests that EMSHMM could better capture the cognitive processes involved in the task and consequently led to higher inference accuracy, demonstrating again the advantage of the EMSHMM approach.

Previous research on inferring participants' preference decisions typically combined eye movement measures with other information such as additional physiological measures or attended visual features. For example, Bee, Prendinger, Nakasone, André, and Ishizuka (2006) integrated skin conductance, blood volume pulse, pupillary response, and eye movement measures to infer participants' preference over two neckties for choice and reached an average accuracy of 81%. Glaholt, Wu, and Reingold (2009) used fixation duration upon different facial features during a face preference decision making task to infer participants' preferences when viewing new faces consisting of old facial feature they saw before and reached about 81% accuracy on average. In contrast to these approaches, here we showed that we could infer participants' preference during a decision making trial from eye gaze transition information alone using EMSHMM with high accuracies (e.g., using the eye movements during the final 2 s before the decision response: Group A, 93%; Group B, 71%; overall average, 81%). This result demonstrated again the power of the EMSHMM approach.

In addition to discovering individual differences in cognitive behavior and inferring participants' preferences, similar to the EMHMM approach, EMSHMM provides quantitative measures of similarities among individual eye movement patterns by calculating the log-likelihood of one's eye movement data being generated by a representative HMM. For example, here we showed that the similarity of participants' eye movement patterns during the preference-biased period to the representative pattern of Group A as opposed to that of Group B (as measured in A–B scale) was positively correlated with the gaze cascade effect and inference accuracy using fixations during the final 2 s before response. In addition to examining the relationship between eye movement patterns and other psychological measures, we may also examine how this eye movement pattern similarity measure would be modulated by factors related to decision making styles, such as gender, cultural, sleep loss, and so forth. Using EMHMM, we previously had shown that eye movement pattern similarity to an eye-centered, analytic pattern during face recognition is associated with better recognition performance whereas similarity to a nose-centered, holistic patterns is correlated with cognitive decline in older adults (e.g., Chan et al., 2018; Chuk, Chan, & Hsiao, 2017; Chuk, Crookes, et al., 2017), and that individuals with insomnia symptoms exhibited eye movement patterns more similar to a representative nose-mouth pattern during facial expression judgments as compared with healthy controls (Zhang, Chan, Lau, & Hsiao, 2019). Similar analyses can be conducted using EMSHMM to examine how eye movement patterns are associated with other psychological measures and factors that may affect eye movement patterns in more complex tasks that involve cognitive state changes.

Note that in the present analysis, we focused on the eye gaze transition behavior between two stimuli of choice in the

preference decision-making task by using only two ROIs, with each corresponding to a stimulus. Future work could extend the analysis to explore the eye movement pattern—that is, the ROIs (low-level states) and transition probabilities among the ROIs—on each stimulus, to capture individual differences in information extraction as well as in gaze transitions in decision-making behavior. Previous studies showed that participants had preferred fixated features or fixation locations during subjective decision-making. For instance, it was found that attractive and unattractive features received more attention than those with intermediate attractiveness (Sutterlin, Brunner, & Opwis, 2008); brands that were located at the center of a shelf in shops were more likely to be chosen (Chandon, Hutchinson, Bradlow, & Young, 2009). Individuals may differ in how they obtain information from the stimuli of choice during decision making (or in a cognitive task in general). In our EMHMM toolbox (Chuk et al., 2014), we capture this individual difference by inferring personalized ROIs on each stimulus using a Gaussian mixture model approach, and the optimal number of ROIs for each participant is estimated through the Bayesian method. This is included as an option in the EMSHMM toolbox presented here.

Note also that here we used a simple decision making task with only two assumed cognitive states to illustrate the EMSHMM method. In this case, if we use a larger number of high-level states, the EMSHMM method will be able to discover more fine-grained cognitive states in between the exploration and preference-biased periods in terms of similarity. In contrast, in more complex cognitive tasks such as driving or cooking, with a large number of high-level states the method will be able to discover more discrete cognitive states essential to the task and their associated eye movement patterns.

In the present study, the SHMM represents differences in transition matrices within a trial (intratrial differences). In contrast, the mixed HMM by Altman (2007) adds random effects to the HMM. In particular, random effects are added to the emission density means and to the log-probabilities of the transition matrix and prior. This allows interparticipant or intertrial differences to be represented in a single model. Thus, future work may extend the SHMM to add random effects to model interparticipant differences in a single model.

In summary, here we proposed a novel HMM-based approach, EMSHMM, to analyze eye movement data in tasks that involve cognitive state changes. For each participant, we used an SHMM to capture between cognitive state transitions during the task, with eye movement patterns during each cognitive state being summarized using a regular HMM. Through applying EMSHMM to a face preference decision-making task, we identified two common eye movement patterns from the participants. One pattern entered the preference-biased cognitive state later and showed a stronger gaze cascade effect immediately before the decision response, and the preference

decision could only be inferred later in a trial. In contrast, the other pattern revealed the preference decision much earlier in a trial and spent more time in the preference-biased cognitive state, but had a weaker gaze cascade effect in the end, leading to a lower decision response inference accuracy. These differences emerged naturally as the result of clustering based on eye movement data alone, and were not revealed by any existing methods in the literature. As compared with our previous EMHMM approach, the EMSHMM method better captured eye movement behavior in the task and thus inferred participants' decision responses with higher accuracy. In addition, EMSHMM provides quantitative measures of similarities among individual eye movement patterns, and thus is particularly suitable for studies using eye movements to examine individual differences in cognitive processes, making a significant impact on the use of eyetracking to study cognitive behavior across disciplines.

Acknowledgments We are grateful to the Research Grant Council of Hong Kong (project 17609117 to J.H.H. and CityU 110513 to A.B.C.) and to JST.CREST (to S.S.). A.B.C. and J.H.H. contributed equally to this article. We thank the editor and two anonymous reviewers for the helpful comments.

Open Practices Statement The code (Matlab Toolbox EMSHMM) and data of the study are available to the research community for noncommercial use at <http://visal.cs.cityu.edu.hk/research/emshmm/>. The experiment reported here was not preregistered.

References

- Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, 36, 623–636.
- Altman, R. (2007). Mixed hidden Markov models: An extension of the hidden Markov model to the longitudinal data setting. *Journal of the American Statistical Association*, 102, 201–210.
- Andrews, T. J., & Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*, 39, 2947–2953. [https://doi.org/10.1016/S0042-6989\(99\)00019-X](https://doi.org/10.1016/S0042-6989(99)00019-X)
- Bee, N., Prendinger, H., Nakasone, A., André, E., & Ishizuka, M. (2006). Auto select: What you want is what you get: Real-time processing of visual attention and affect. In *International Tutorial and Research Workshop on Perception and Interactive Technologies for Speech-Based Systems* (pp. 40–52). Heidelberg, Germany: Springer.
- Camci, F., & Chinnam, R. B. (2006). Hierarchical HMMs for autonomous diagnostics and prognostics. In *Proceedings of the 2006 IEEE International Joint Conference on Neural Networks* (pp. 2445–2452). Piscataway, NJ: IEEE Press.
- Castelhano, M. S., & Henderson, J. M. (2008). Stable individual differences across images in human saccadic eye movements. *Canadian Journal of Experimental Psychology*, 62, 1–14. <https://doi.org/10.1037/1196-1961.62.1.1>
- Chan, C. Y. H., Chan, A. B., Lee, T. M. C., & Hsiao, J. H. (2018). Eye movement patterns in face recognition are associated with cognitive decline in older adults. *Psychonomic Bulletin & Review*, 25, 2200–2207.
- Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing*, 73(6), 1–17. <https://doi.org/10.1509/jmkg.73.6.1>
- Chuk, T., Chan, A. B., & Hsiao, J. H. (2014). Understanding eye movements in face recognition using hidden Markov models. *Journal of Vision*, 14(11), 8:1–14. <https://doi.org/10.1167/14.11.8>
- Chuk, T., Chan, A. B., & Hsiao, J. H. (2017). Is having similar eye movement patterns during face learning and recognition beneficial for recognition performance? Evidence from hidden Markov modeling. *Vision Research*, 141, 204–216.
- 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.
- Coutrot, A., Hsiao, J. H., & Chan, A. B. (2018). Scanpath modeling and classification with hidden Markov models. *Behavior Research Methods*, 50, 362–379. <https://doi.org/10.3758/s13428-017-0876-8>
- Coviello, E., Chan, A. B., & Lanckriet, G. R. (2014). Clustering hidden Markov models with variational HEM. *Journal of Machine Learning Research*, 15, 697–747.
- Ferrari, J. R., & Dovidio, J. F. (2000). Examining behavioral processes in indecision: Decisional procrastination and decision-making style. *Journal of Research in Personality*, 34, 127–137.
- Germeijs, V., & De Boeck, P. (2002). A measurement scale for indecisiveness and its relationship to career indecision and other types of indecision. *European Journal of Psychological Assessment*, 18, 113–122.
- Glaholt, M. G., Wu, M. C., & Reingold, E. M. (2009). Predicting preference from fixations. *Psychology Journal*, 7, 141–158.
- Haji-Abolhassani, A., & Clark, J. J. (2014). An inverse Yarbus process: Predicting observers' task from eye movement patterns. *Vision Research*, 103, 127–142.
- Hariri, B., Shirmohammadi, S., & Pakravan, M. R. (2008). A hierarchical HMM model for online gaming traffic patterns. In *Proceedings of 2008 IEEE Instrumentation and Measurement Technology Conference* (pp. 2195–2200). Piscataway, NJ: IEEE Press.
- Hayhoe, M., & Ballard, D. (2014). Modeling task control of eye movements. *Current Biology*, 24, R622–R628.
- Henderson, J. M., Shinkareva, S. V., Wang, J., Luke, S. G., & Olejarczyk, J. (2013). Predicting cognitive state from eye movements. *PLoS ONE*, 8, e64937. <https://doi.org/10.1371/journal.pone.0064937>
- Kanan, C., Bseiso, D. N., Ray, N. A., Hsiao, J. H., & Cottrell, G. W. (2015). Humans have idiosyncratic and task-specific scanpaths for judging faces. *Vision Research*, 108, 67–76.
- Lemonnier, S., Brémond, R., & Baccino, T. (2014). Discriminating cognitive processes with eye movements in a decision-making driving task. *Journal of Eye Movement Research*, 7(4), 3:1–14.
- Liechty, J., Pieters, R., & Wedel, M. (2003). Global and local covert visual attention: Evidence from a Bayesian hidden Markov model. *Psychometrika*, 68, 519–541.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (pp. 281–297). Berkeley, CA: University of California Press.
- Poynter, W., Barber, M., Inman, J., & Wiggins, C. (2013). Individuals exhibit idiosyncratic eye-movement behavior profiles across tasks. *Vision Research*, 89, 32–38.
- Rassin, E., Muris, P., Booster, E., & Kolsloot, I. (2008). Indecisiveness and informational tunnel vision. *Personality and Individual Differences*, 45, 96–102.
- Risko, E. F., Anderson, N. C., Lanthier, S., & Kingstone, A. (2012). Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. *Cognition*, 122, 86–90.

- Sekiguchi, T. (2011). Individual differences in face memory and eye fixation patterns during face learning. *Acta Psychologica*, 137, 1–9. <https://doi.org/10.1016/j.actpsy.2011.01.014>
- Shimojo, S., Simion, C., & Changizi, M. A. (2011). Gaze and preference-orienting behavior as a somatic precursor of preference decision. In R. B. Adams (Ed.), *The science of social vision* (pp. 151–163). Oxford, UK: Oxford University Press.
- Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, 6, 1317–1322.
- Simion, C., & Shimojo, S. (2007). Interrupting the cascade: Orienting contributes to decision making even in the absence of visual stimulation. *Perception & Psychophysics*, 69, 591–595. <https://doi.org/10.3758/BF03193916>
- Simola, J., Salojärvi, J., & Kojo, I. (2008). Using hidden Markov model to uncover processing states from eye movements in information search tasks. *Cognitive Systems Research*, 9, 237–251.
- Sutterlin, B., Brunner, T. A., & Opwis, K. (2008). Eye-tracking the cancellation and focus model for preference judgments. *Journal of Experimental Social Psychology*, 44, 904–911.
- Van der Lans, R., Pieters, R., & Wedel, M. (2008). Eye-movement analysis of search effectiveness. *Journal of the American Statistical Association*, 103, 452–461.
- Wu, D. W. L., Bischof, W. F., Anderson, N. C., Jakobsen, T., & Kingstone, A. (2014). The influence of personality on social attention. *Personality and Individual Differences*, 60, 25–29.
- Yeung, P. Y., Wong, L. L., Chan, C. C., Leung, J. L., & Yung, C. Y. (2014). A validation study of the Hong Kong version of Montreal Cognitive Assessment (HK-MoCA) in Chinese older adults in Hong Kong. *Hong Kong Medical Journal*, 20, 504–510.
- Yi, W., & Ballard, D. H. (2009). Recognizing behavior in hand-eye coordination patterns. *International Journal of Humanoid Robotics*, 6, 337–359.
- Zhang, J., Chan, A. B., Lau, E. Y. Y., & Hsiao, J. H. (2019). Individuals with insomnia misrecognize angry faces as fearful faces while missing the eyes: An eye-tracking study. *Sleep*, 42, zsy220. <https://doi.org/10.1093/sleep/zsy220>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.