Is having similar eye movement patterns during face learning and recognition beneficial for recognition performance? Evidence from hidden Markov modeling

Tim Chuk[1], Antoni B. Chan[2], Janet Hsiao[1]

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

(This is a draft version of a paper accepted for publication in the journal Vision Research.)

Address for correspondence:
Janet Hsiao
Department of Psychology
University of Hong Kong
PokFulam Road Hong Kong
Phone: +852 3917 4874
Email: jhsiao@hku.hk

Abstract

The hidden Markov model (HMM)-based approach for eye movement analysis is able to reflect individual differences in both spatial and temporal aspects of eye movements. Here we used this approach to understand the relationship between eye movements during face learning and recognition, and its association with recognition performance. We discovered holistic (i.e., mainly looking at the face center) and analytic (i.e., specifically looking at the two eyes in addition to the face center) patterns during both learning and recognition. Although for both learning and recognition, participants who adopted analytic patterns had better recognition performance than those with holistic patterns, a significant positive correlation between the likelihood of participants' patterns being classified as analytic and their recognition performance was only observed during recognition. Significantly more participants adopted holistic patterns during learning than recognition. Interestingly, about 40% of the participants used different patterns between learning and recognition, and among them 90% switched their patterns from holistic at learning to analytic at recognition. 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, participants who used the same or different patterns during learning and recognition did not differ in recognition performance. The similarity between their learning and recognition eye movement patterns also did not correlate with their recognition performance. These findings suggested that perceptuomotor memory elicited by eye movement patterns during learning does not play an important role in recognition. In contrast, the retrieval of diagnostic information for recognition, such as the eyes for face recognition, is a better predictor for recognition performance.

Keywords: individual difference, eye movement, hidden Markov model, face learning, face recognition

Introduction

In human vision, the density of photoreceptors on the retina is not uniform. It is extremely high at the fovea, and drops dramatically as visual eccentricity increases. Thus, the fovea has the highest visual acuity, whereas the perifoveal area, which is much larger than the fovea, is of low visual acuity. In order for an individual to see clearly a region of interest in a cognitive task, the fovea has to be constantly relocated to the region (Tovee, 1996). Consequently, our eyes are constantly moving, and eye movements are shown to reflect underlying cognitive processes, or more specifically the way information is sampled from the environment (Antrobus, Antrobus, & Singer, 1964; Yarbus, 1967; Grant & Spivey, 2003; Heremans, Helsen, &Feys, 2008). Thus, it is reasonable to speculate that different eye movement patterns may lead to different performances in cognitive tasks.

Consistent with this speculation, it has been reported that in a cognitive task, experts and novices typically exhibited different eye movement patterns. For instance, Charness et al. (2001) reported that expert and intermediate chess players have different eye movement patterns. Experts made significantly more fixations at empty squares on the board. They also fixated significantly more often at pieces relevant to the current task than did the intermediates. Waters and Underwood (1998) compared the eye movement patterns of expert and novice musicians when they participated in a simple music reading task. The participants were shown two melodic fragments successively, and asked to judge whether the two fragments were the same or different. It was found that experts made significantly more fixations at the first fragment than novices and that their initial fixations were of significantly shorter duration than the novices. Similar findings were also reported in the research on reading. Siyanova-Chanturia, Conklin, and Schmitt (2011) compared the eye movement patterns of native and non-native English speakers when they were asked to read idioms and novel phrases. It was found that native speakers made significantly fewer and shorter fixations at idioms than novel phrases. In contrast, the number and duration of fixations that non-native speakers made at idioms and novel phrases were similar to each other. This demonstrated that native speakers had a processing advantage for idioms over novel phrases, which was not presented among non-native speakers. Hyona, Lorch, &Kaakinen (2002) compared eye movement patterns of native Finnish speakers when they were reading Finnish texts and found that those who fixated more often at the headings and topic-final sentences performed significantly better than those who showed other eye movement patterns when they were required to summarize the texts.

Nevertheless, in the literature on face recognition, it remains controversial whether different eye movement patterns are associated with different recognition performances. For example, Goldinger, He, and Papesh (2009) found that in a face recognition memory task, participants made fewer fixations, visited fewer regions of interest, and had shorter scanning distances on the trials in which they failed to recognize a learned face as compared with those that led to successful recognition. Glen et al. (2012) found that among people who suffered from central visual field defects, those who performed better in face recognition demonstrated a different eye movement strategy as compared with the ones who performed worse. These findings suggest that eye movement patterns are associated with performance in face recognition. In contrast, Blais et al. (2008) found that in face recognition, although Asian participants looked primarily at the center of the faces (i.e., a holistic scanning pattern) whereas Caucasian participants looked more frequently at facial features such as the two eyes and the mouth (i.e., an analytic pattern), the two cultural groups showed comparable recognition performance. This finding was later replicated in Caldara, Zhou, and Miellet (2010). Similarly, Mehoudar, Arizpe, Baker, & Yovel (2014) found that participants showed idiosyncratic eye movement patterns in face recognition that were highly stable over time; however, these patterns were not predictive of their recognition performance.

These inconsistent findings in the literature may be due to substantial individual differences in eye movement pattern that were not adequately reflected in the data analyses. Indeed, recent studies have shown that there are considerable individual differences in eye movement that persist over time and across different stimuli when people perform cognitive tasks. For instance, Castelhano and Henderson (2008) showed that during picture viewing, the characteristics of fixation durations and saccade amplitudes in eye movement differed across individuals but were stable within an individual across different types of visual stimuli. Risko et al. (2012) found that curiosity was a significant predictor of participants' eye movement patterns in scene viewing. Peterson and Eckstein (2013) showed that participants differed significantly in where to first move their eyes in a face identification task, and they performed better when being forced to look at their preferred viewing locations than other locations. Kanan, Bseiso, Ray, Hsiao, and Cottrell (2015) showed that the identity of participants could be inferred based on their eye movements across different face perception judgment tasks. These findings provided stronger evidence for the existence of substantial individual differences in eye movement.

In order to account for individual differences in both spatial (i.e., fixation locations) and temporal dimensions (i.e., transitions among fixation locations) of eye movement in the data analysis, in our previous study (Chuk, Chan, & Hsiao, 2014a), we proposed to use a hidden Markov model (HMM) to summarize an individual's eye movement pattern in face recognition. The hidden states of the HMM represented the individual's regions of interests (ROIs) for eye fixations. The individual's eye movements among the ROIs were summarized through the HMM's transition matrix, which represents the probability of each ROI being viewed next conditioned on the currently viewed ROI. The process of learning the individual HMMs was completely data driven. The individual HMMs could then be clustered based on their similarities to discover common patterns shared by individuals. The similarity of an individual pattern to a common pattern discovered through clustering could be measured as the likelihood of the individual pattern being classified as the common pattern. Through this approach, we discovered two common eye movement patterns in face recognition within our Asian participants that resembled the holistic and analytic patterns found in Asian and Caucasian participants respectively in Blais et al. (2008) and Caldara et al. (2010). This finding showed that both eye movement patterns could be observed within a cultural group, demonstrating substantial individual differences in eye movement pattern. In our follow-up study (Chuk et al., 2014b; Chuk, Crookes, Hayward, Chan, & Hsiao, submitted), we found that analytic and holistic patterns could be observed in both Asians and Caucasians, and the two cultural groups did not differ significantly in the percentage of group members being classified as using holistic or analytic patterns. Also, the participants who showed analytic eye movement patterns performed significantly better than those who showed holistic patterns, and there was a positive correlation between the likelihood of participants' pattern being classified as analytic and their recognition performance. These findings were not possible without taking individual differences in eye movement into account, demonstrating well the advantage of our HMM approach.

Our results from previous studies suggested that analytic eye movement patterns, which involved eye fixations specifically to the two eyes in addition to the face center, were beneficial for face recognition. This result was consistent with the previous studies showing that the eyes are the most important features for face recognition (e.g., Gosselin & Schyns, 2001; Vinette, Gosselin, & Schyns, 2004). For example, using the Bubbles technique, Gosselin and Schyns (2001) found that the two eyes were the most diagnostic features for recognizing the identity of an individual. Vinette et al. (2004) further showed that the left eye was the earliest diagnostic feature that participants used in face recognition. Afterwards, both the left and right eyes were used effectively.

Nevertheless, it remains unclear whether analytic eye movement patterns are also beneficial for face learning. Henderson, William, and Falk (2005) found that when participants' eye movements were restricted to be at the face center during the learning phase of a face recognition task, their performance in the recognition phase was impaired significantly. This result suggested that the eye movements during the learning phase were related to recognition performance. Sekiguchi (2011) further showed that participants who had high face recognition memory performance moved their eyes between the left and right eyes more frequently (i.e., an analytic eye movement pattern) during face learning than those with low recognition performance. This result suggests that, similar to eye movements during face recognition, analytic eye movement patterns during face learning may also be associated with better recognition performance.

In addition, in the literature, it has been suggested that during visual recognition, participants showed similar eye movements to those generated during visual learning. For instance, the scan path theory posits that in pattern perception, the mental representation of visual patterns includes the perceptuomotor cycle involved during memory encoding. Accordingly, eye movements produced during learning have to be repeated during recognition for the recognition to be successful (Noton & Stark, 1971a; 1971b). Consistent with this theory, Laeng and Teodorescu (2002) found that when participants were asked to recall a learned picture in front of a whiteboard, they had better performance when their eyes were allowed to move freely than restricted to be at the center of the board, and their eye movements resembled those generated during learning. In face recognition, Blais et al. (2008) found that although in general, participants in the recognition phase made fewer fixations than in the learning phase, their eye movements did not show any significant difference in terms of fixation location or duration during the two phases. More specifically, Asian participants consistently showed holistic eye movement patterns whereas Caucasian participants showed analytic patterns in both the learning and recognition phases (see also Caldara et al., 2010). In contrast, some studies have shown that an exact repetition of eye movements during learning was not necessary for successful recognition. For example, participants were able to recognize previously learned visual stimuli in tachistocscopic presentations, in which eye movements were not possible (e.g., Thorpe, Fize, & Marlot, 1996). They were also able to recognize faces when their eye gaze was restricted to be at the face center during learning, and their eye movements during recognition were similar to those generated when they were allowed to move their eyes freely during learning (Henderson et al., 2005). However, these results did not completely rule out the influence of perceptuomotor memory in pattern recognition as suggested in the scan path theory. It remains possible that participants who show more similar eye movement patterns during face learning and recognition perform better in face recognition than those who show different patterns.

Indeed, eye movements during pattern recognition can be influenced by multiple factors in addition to perceptuomotor memory, such as top-down expectations and bottom-up image saliency, and thus eye movement patterns during recognition may not be exact replications of those generated during learning (e.g., Henderson, 2003; Rayner, 1998; Yarbus, 1965). Accordingly, eye movements during learning and recognition should differ because the two phases involve different task expectations and cognitive processes: information encoding during learning, and information retrieval during recognition. Consistent with this speculation, Hsiao and Cottrell (2008) found that during face learning and recognition participants showed different fixation duration profiles: during face learning, participants' first fixations were short, and the duration gradually increased for the second and then the third fixations, whereas during recognition, there was no difference between the first three fixations in terms of duration. Nevertheless, in contrast to this finding, Blais et al. (2008) reported that participants' eye movement patterns during face learning and recognition did not differ significantly in either fixation location or duration (see also Caldara, et al., 2010). We speculate that this inconsistency may be because participants differed in whether they used similar eye movement strategies for face learning and recognition, and this individual difference might have been obscured in group-level analysis used in previous studies. In addition, this individual difference may also be related to their recognition performance, as suggested by the scan path theory. Such examination requires individual-level eye movement pattern analysis.

Thus, here we aimed to examine whether participants used different eye movement patterns for face learning and recognition through individual-level data analysis using the HMM based approach. We also aimed to examine whether eye movement patterns during face learning were associated with performance during the recognition phase, and whether the similarities between participants' eye movement patterns during face learning and recognition were related to their recognition performance. In view of the previous finding that eye movements in face learning and recognition may differ in fixation duration (Hsiao & Cottrell, 2008), in the current study we included fixation duration information in addition to fixation location information in the HMMs. This expansion of the model allowed us to model participants' eye movement patterns more precisely. We hypothesized that: 1) During face learning, common eye movement patterns similar to the holistic and analytic patterns discovered during face recognition may also be observed, and participants with analytic patterns during face learning may also perform better in face recognition than those with holistic patterns; 2) Participants may use different eye movement patterns during face learning and recognition, reflecting different underlying cognitive processes; 3) Individuals differ in the similarity between eye movement patterns during face learning and recognition, and this similarity may be associated with their recognition performance, as suggested by the scan path theory.

Method

Behavioral task

Here we used the data collected in Chuk, Crookes, Hayward, Chan, and Hsiao (submitted; see also Chuk et al., 2014b) for the data analysis. A total of 48 participants (24 Asians and 24 Caucasians) were recruited for a face recognition task. The mean age of Asian participants (7 males) was 21.5 (SD = 2.2), whereas that of Caucasian participants (6 males) was 21.2 (SD = 7.5). The task had two sessions, one with Asian face images and the other with Caucasian face images (counterbalanced across participants). Each session had a learning phase and a recognition phase. There was no time delay between the two phases, but participants were allowed to take a break between the two sessions. Participants in each learning phase were required to view 14 faces one at a time, each for 5 seconds. In each recognition phase, they were presented with the 14 learned faces and 14 new faces one at a time, and were required to judge through button responses whether they saw the face during the learning phase or not; the face image stayed on the screen until the response. During both phases, participants started each trial with a central fixation cross. The face image was then presented at one of the four quarters on the screen in a random order. The distance between the central fixation cross and the image locations subtended about 9 degrees of visual angle horizontally and about 7 degrees vertically. The face images subtended about 8 degrees of visual angle horizontally and 13 degrees vertically. Participants' eye movements were recorded with an EyeLink 1000 eye tracker.

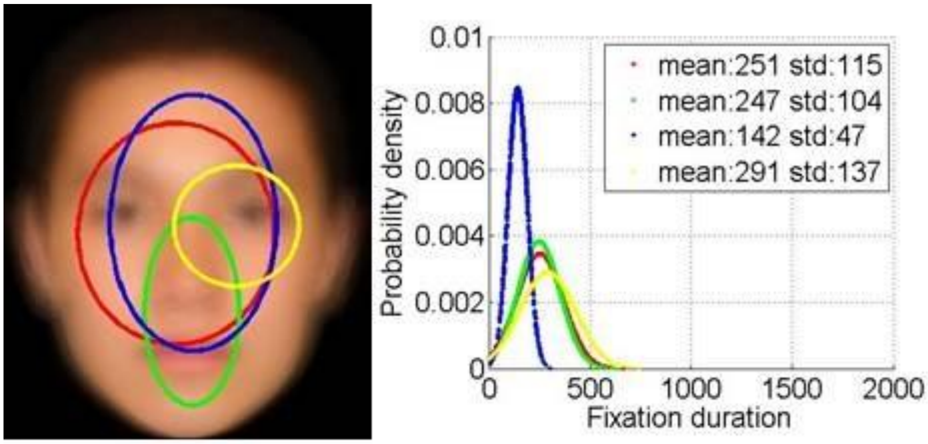
Eye movement data were extracted from the EyeLink 1000 system using the default software Data Viewer. In data acquisition, the EyeLink 1000 defaults for cognitive research was used: saccade motion threshold was 0.15 degree of visual angle; saccade acceleration threshold was 8000 degree / square second; saccade velocity threshold was 30 degree / second. The software produced a fixation report for each participant. We then filtered out fixations that were not located in the face area. The remaining eye movement data were used for data analyses.

Hidden Markov models

We assumed that a participant's eye movements in a cognitive task could be summarized with a hidden Markov model (HMM), so that we were able to examine individual differences in eye movements through comparing individual HMMs. Furthermore, we clustered the individual HMMs to discover common patterns (the toolbox, Eye movement hidden Markov models approach (EMHMM), can be downloaded here: http://visal.cs.cityu.edu.hk/research/emhmm/).

HMMs are a type of time-series model that assumes that the observed time-series data arise from an underlying state process, where the current state depends only on the previous state. The underlying states are hidden; they can be estimated from the probabilistic association between the observed data and the states (i.e., the emission density of a state), as well as from the transition probabilities between the states. An HMM contains a vector of prior values, which indicates the probability of a time-series beginning with each state; a transition matrix, which specifies the transition probabilities between any two hidden states; and a Gaussian emission for each state, which represents the probabilistic association between the observed data (e.g., eye fixation locations) and a hidden state.

In the context of eye movement analysis here, the observed time series were eye fixation sequences, with each observation consisting of both fixation location and fixation duration. Each hidden state of the HMM represented a Region of Interest with Duration (ROID), which contained the location of the region of interest (ROI), as well as fixation duration in the ROI (Note that in our earlier implementation reported in Chuk et al., 2014, we did not include duration information). We assumed that both the locations and durations of the fixations belonging to an ROID followed a Gaussian distribution (see Ohl, Brandt, & Kliegl, 2013). Each ROID therefore was represented as a three-dimensional Gaussian emission, where two dimensions corresponded to the spatial distributions of the fixations (i.e., fixation locations), and the third dimension corresponded to the temporal distribution of the fixations (i.e., fixation durations). In the HMM, the prior vector indicated the probabilities that a fixation sequence started in a particular ROID, while the transition matrix contained the probabilities of moving to the next ROID from the current ROID. Figure 1 shows an example HMM. An acyclic graph whose nodes represent the components of the HMM is shown in Figure 2.



	to red	to green	to blue	to yellow
priors	.35	.03	.57	.05
from red	.61	.09	.02	.28
from green	.40	.22	.13	.25
from blue	.53	.06	.15	.26
from yellow	.61	.22	.03	.14

Figure 1. An example of an HMM summarizing eye fixation data. The ellipses on the face represent the location of the ROIDs. The ellipse represents 2 standard deviations around the mean of the Gaussian spatial distribution. The one-dimensional Gaussian distributions on the right show the fixation durations of the corresponding ROID. The table presents the transition probabilities between the ROIDs. Note that the red and blue ROIDs are spatially overlapping, but have different fixation durations.

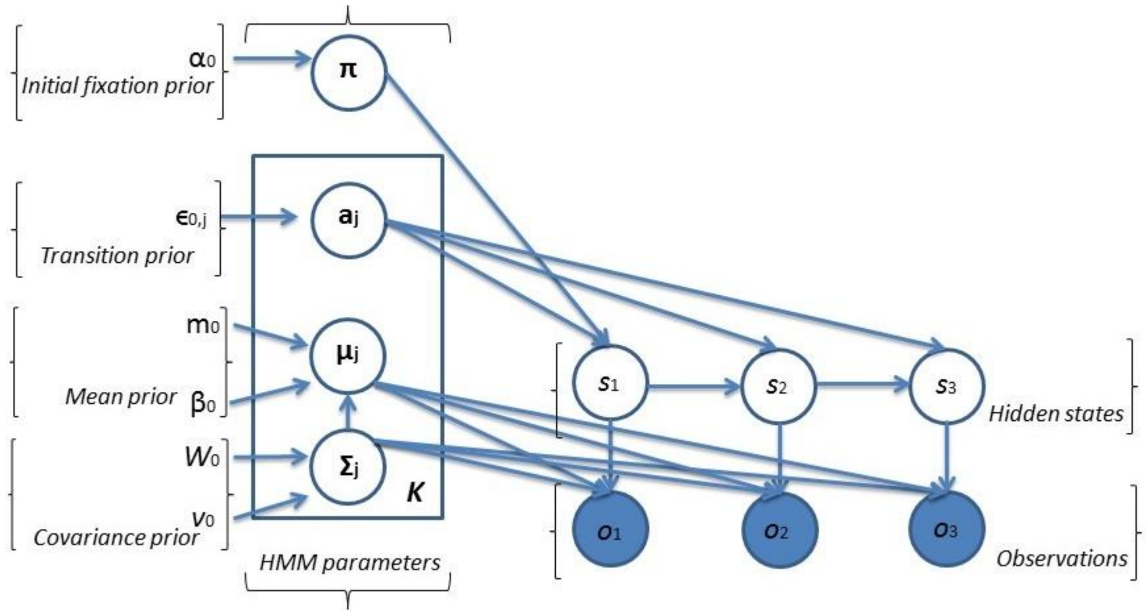


Figure 2. An acyclic graph representing the components, the parameters, and the hyperparameters of an HMM. The colored nodes (o_n) represent the observed fixation data; the nodes on top (s_n) represent the hidden states. The nodes on the left represent the prior distributions of the HMM parameters; K represents the number of hidden states (ROIDs), which is determined by the algorithm. The symbols left to the nodes represent the hyper-parameters of the prior distributions.

For each participant, we trained two HMMs using either the eye movement data from the learning phase (learning phase HMM) or the data from the recognition phase (recognition phase HMM). We implemented the variational Bayesian expectation-maximization (VBEM) algorithm (Bishop, 2006) in Matlab to estimate the parameters of the HMMs. This Bayesian approach places a prior distribution on each parameter of the model and then approximates the posterior distribution of the parameters using a factorized variational distribution. The prior distributions for the Gaussian emissions were Normal-Wishart distributions. For the spatial dimensions, we set the prior mean to be the center of the image (m_0 in Figure 2). The covariance matrices of the Gaussians were set to be isotropic matrices with standard deviation of 14 pixels (0.53 degree of visual angle) for the spatial dimensions (W_0 in Figure 2), which was about the same size as a facial feature on the image. For the temporal dimensions, we set the prior mean and prior standard deviation using the fixation durations at the population level. The hyper-parameter v_0 for the covariance matrices was set to 5, and the hyper-parameter β_0 for the means was set to 1. The prior distributions for the transition matrix and prior vector were Dirichlet distributions, and we set the concentration parameter to 0.005 to reflect the assumption that the number of ROIDs on a face was much fewer than the number of fixation locations.

The VBEM algorithm for estimating an HMM proceeded as follows. First, we initialized the transition matrices (ε_0 in Figure 2) and prior vectors (α_0 in Figure 2) as uniform distributions, and we obtained the initial Gaussian emissions (ROIDs) using the Matlab “fit” function for Gaussian mixture models. The VBEM algorithm then iterates between the E-step and the M-step until convergence. In the E-step, the forward-backward algorithm is used to calculate the single and pairwise responsibilities, corresponding to the marginal probability of a state at a particular time and the joint probability of two consecutive states, respectively. In the M-step, we updated the model parameters using the calculated responsibilities. All parameters of the HMMs were updated simultaneously during the E-M loop. To avoid convergence to a local maximum, we trained the model 100 times with different initial Gaussian ROIDs calculated by the Matlab fit function, and selected the model with the highest log-likelihood of the data.

Finally, for each individual, we determined the number of hidden states (ROIDs) in their HMM in a data-driven fashion. In our previous study (Chuk et al., 2014), the number of hidden states for each model (HMM) was set to 3. In the current study, we implemented automatic model selection. We trained six separate HMMs with different numbers of ROIDs, ranging from 1 to 6. We then selected the HMM from this set with the highest log-likelihood of the data, thus determining the number of ROIDs for the individual. On our data, this selection method typically selected three or four hidden states. Note that we used a Bayesian methodology that automatically penalizes model complexity via the prior distributions on the model parameters. Hence, the selected model was the most parsimonious explanation of the data.

Clustering Hidden Markov models

In order to discover the common fixation patterns shared by participants, we clustered the individuals' HMMs into groups using the hierarchical variational expectation maximization (VHEM) algorithm (Coviello et al., 2012). For each group, VHEM generates a representative HMM that describes the ROIDs and transition probabilities for the common pattern used by the group. Furthermore, the log-likelihood of each participant's eye movement data was calculated with respect to each representative HMM, which yielded a measure of how similar their eye movement patterns were to the common patterns. For each participant and representative HMM, we calculated the average of the log-likelihoods of the fixation sequences over all trials. For each trial, the log-likelihood was normalized by dividing by the length of the sequence, in order to remove the effect of different sequence lengths (Oates, Firoui & Cohen, 2001; Seo, Kishino, & Thorne 2005; Martin, Hurn, & Harris, 2012). This measure was correlated with the participant's recognition performance in order to reveal whether certain common patterns were associated with better performance.

We applied the above clustering method separately for the learning phase and the recognition phase HMMs: the 48 learning phase HMMs were clustered into groups, and the 48 recognition phase HMMs were also clustered into groups. We clustered the participants' HMMs into two groups for each phase because several previous studies (e.g., Blais et al., 2008; Kelly et al., 2011; Chuk, Chan, & Hsiao, 2014a) showed that most people's eye movement patterns exhibited one of the two fixation patterns: a holistic pattern that focused mainly at the center of the face, or an analytic pattern that focused at specific facial features (e.g., the two eyes and the mouth) in addition to the face center. Since we used a variational Bayesian approach to estimate parameters of individual HMMs, the input HMMs may have different numbers of hidden states. In the current modeling, the majority of the individual HMMs ended up having four ROIDs, and thus we set the representative HMMs in the VHEM algorithm to have four hidden states.

Previous studies (e.g. Hsiao & Cottrell, 2008) showed that participants had different eye movement patterns during the learning and the recognition phases, and that this difference was at least partly in terms of fixation duration. Therefore, we also tried to cluster the individuals' learning and recognition phase HMMs together into two clusters to see if participants indeed changed their eye movement strategies during the two phases and whether the change was related to their recognition performance.

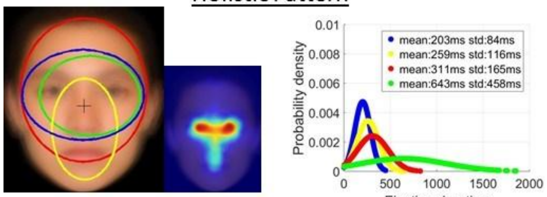
Results

Eye movement patterns during the learning phase

To discover common eye movement patterns participants used during the learning phase, we modeled each participant's eye movements during the learning phase with an HMM and clustered the individual HMMs into two groups. Figure 3a shows the representative HMMs of the two resulting groups. Table 1 shows the number of participants being clustered into each eye movement pattern group.

(a) Learning Phase

Holistic Pattern



	To Blue	To Yellow	To Red	To Green
Priors	.62	.14	.21	.03
From Blue	.56	.11	.28	.05
From Yellow	.28	.47	.22	.03
From Red	.30	.09	.55	.06
From Green	.32	.15	.15	.38

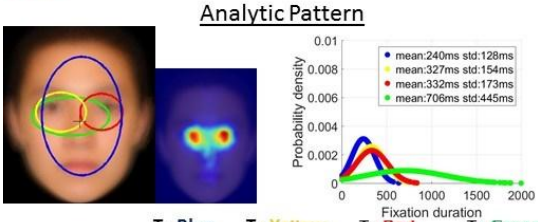
Mean location (relative to the face center) and standard deviation of ROIDs in visual angle

	Mean X	Mean Y	SD X	SD Y
Blue	-.04	-.40	3.47	2.69
Yellow	.02	.85	1.85	3.03
Red	.00	-.59	3.59	4.28
Green	.21	-.38	2.87	2.39

Mean location (relative to the face center) and standard deviation of ROIDs in face-size-normalized unit

	Mean X	Mean Y	SD X	SD Y
Blue	-.01	-.06	.76	.44
Yellow	.00	.14	.41	.48
Red	.00	-.09	.79	.69
Green	.05	-.06	.63	.39

Analytic Pattern



	To Blue	To Yellow	To Red	To Green
Priors	.89	.04	.05	.02
From Blue	.52	.14	.27	.07
From Yellow	.38	.23	.37	.02
From Red	.44	.39	.16	.01
From Green	.44	.05	.02	.49

Mean location (relative to the face center) and standard deviation of ROIDs in visual angle

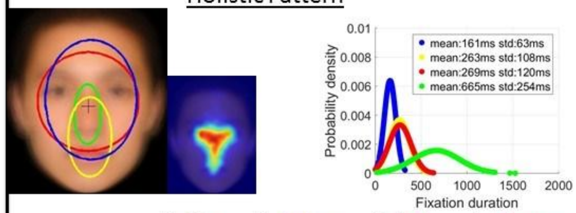
	Mean X	Mean Y	SD X	SD Y
Blue	.08	-.21	2.49	3.75
Yellow	-.72	-.36	1.59	1.25
Red	.83	-.36	1.32	1.14
Green	-.34	-.19	2.42	1.21

Mean location (relative to the face center) and standard deviation of ROIDs in face-size-normalized unit

	Mean X	Mean Y	SD X	SD Y
Blue	.02	-.03	.55	.59
Yellow	-.16	-.06	.35	.19
Red	.18	-.06	.30	.18
Green	-.07	-.03	.53	.19

(b) Recognition Phase

Holistic Pattern



	To Blue	To Yellow	To Red	To Green
Priors	.77	.02	.21	< .01
From Blue	.29	.12	.58	.01
From Yellow	.19	.25	.56	< .01
From Red	.06	.12	.82	< .01
From Green	.66	< .01	.20	.14

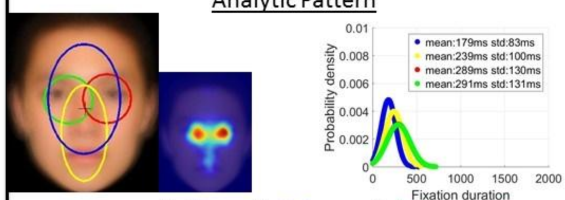
Mean location (relative to the face center) and standard deviation of ROIDs in visual angle

	Mean X	Mean Y	SD X	SD Y
Blue	.09	-.25	2.64	3.56
Yellow	.06	1.12	1.25	2.35
Red	.00	-.17	2.95	2.96
Green	-.02	.28	.76	1.78

Mean location (relative to the face center) and standard deviation of ROIDs in face-size-normalized unit

	Mean X	Mean Y	SD X	SD Y
Blue	.02	-.04	.58	.57
Yellow	.01	.18	.28	.38
Red	.00	-.03	.64	.47
Green	.00	.05	.17	.28

Analytic Pattern



	To Blue	To Yellow	To Red	To Green
Priors	.85	.04	.07	.04
From Blue	.30	.04	.39	.27
From Yellow	.25	.32	.27	.16
From Red	.28	.27	.15	.30
From Green	.22	.10	.58	.10

Mean location (relative to the face center) and standard deviation of ROIDs in visual angle

	Mean X	Mean Y	SD X	SD Y
Blue	-.04	-.36	2.19	3.34
Yellow	-.04	.91	1.40	2.92
Red	.81	-.38	1.47	1.48
Green	-.62	-.32	1.51	1.44

Mean location (relative to the face center) and standard deviation of ROIDs in face-size-normalized unit

	Mean X	Mean Y	SD X	SD Y
Blue	-.01	-.06	.48	.53
Yellow	-.01	.15	.30	.47
Red	.18	-.06	.31	.24
Green	-.14	-.05	.33	.23

Figure 3. The representative HMMs of the two common eye movement patterns discovered by clustering the HMMs for (a) the learning phase, and (b) the recognition phase. The figure shows the spatial distribution of the ROIDs and the corresponding heat

map, the duration distribution of the ROIDs (in ms), and the transition probability matrix of the ROIDs. The tables below the transition matrix show the mean location (relative to the face center) and standard deviation of the ROIDs in visual angle and in face-size-normalized unit. Note that in (b), in the analytic pattern during the recognition phase, the red and green ROIDs had very similar duration distributions, and thus the curves are overlapped.

Learning phase			
(a)	Female	Male	**Total**
Holistic pattern	25	9	**34**
Analytic pattern	10	4	**14**
(b)	Caucasian	Asian	
Holistic pattern	16	18	**34**
Analytic pattern	8	6	**14**
Recognition phase			
(a)	Female	Male	**Total**
Holistic pattern	16	4	**20**
Analytic pattern	19	9	**28**
(b)	Caucasian	Asian	
Holistic pattern	11	9	**20**
Analytic pattern	13	15	**28**

Table 1. The number of participants being clustered into each eye movement pattern group (analytic vs. holistic) with a breakdown by gender (a) and by race (b), using the representative HMMs in Figure 3.

It can be seen that in the holistic representative HMM in Figure 3a, three of the four ROIDs (except the yellow ROID) were centered at the bridge of the nose (i.e., the center of the face). Participants in this group typically started a trial by looking at the center of the face with a short fixation (M = 203 ms, blue ROID). Afterwards, they most likely remained looking at the center with either a short fixation, (M = 203 ms) or a long fixation (M = 311 ms, about 28% of the times, red ROID), and sometimes (11%) looked at the tip of the nose/mouth region (duration M = 259 ms, yellow ROID). Occasionally (about 5%) they made a very long fixation (M = 643 ms, green ROID) at the center of the face. Since in this pattern, participants mainly looked at the center of the face, we refer to this pattern as the holistic pattern during the learning phase.

In the analytic representative HMM shown in Figure 3a, the blue ROID was at the center of the face, whereas a smaller, green ROID was slightly to the left of the center, between the left eye and the bridge of the nose. The yellow and red ROIDs were located at the left and right eye respectively. Participants in this group typically started a trial by looking at the center of the face with a short fixation (M = 240 ms, blue ROID). Afterwards, they either remained looking at the center of the face with short fixations (blue ROID) or started looking at the two eyes (yellow and red ROIDs). When they looked at the two eyes, the fixations were all with long duration (the left eye, M = 327 ms; the right eye, M = 332 ms). Occasionally (6%), they looked between the left eye and the bridge of the nose with a long fixation (M = 706 ms, green ROID). Since in this pattern, participants looked at the two eyes specifically in addition to the face center, we refer to this pattern as the analytic pattern during the learning phase.

The two patterns showed a few similarities and differences. For both patterns, there was an ROID with longer mean fixation duration (M > 600 ms) than the other ROIDs, centered around the bridge of the nose (i.e., center of the face). However, the analytic pattern had two ROIDs on the two eyes with relatively long fixation durations (M > 300 ms), which suggested that participants in this group looked specifically at the two eyes with long fixation durations. In contrast, in the holistic pattern, the ROIDs were mostly at the center of the face. These results suggested that people who showed holistic patterns did not look at the eyes as much and as long as those who showed analytic patterns. There were in total 34 participants who showed holistic patterns during the learning phase; the other 14 participants showed analytic patterns. There were significantly more participants showing holistic patterns than analytic patterns, χ^2 (1) = 8.33, *p*= .003 (Table 1, learning phase).

We then compared the recognition performance of participants showing different eye movement patterns during the learning phase using A-prime (A’). A-prime is a non-parametric alternative to d-prime (d’) and thus can be estimated when the hit or the false-alarm rate was zero. The equations for A’ are shown below.

$$A' = \begin{cases} .5 + \dfrac{(\mathrm{H}-\mathrm{F})(1+\mathrm{H}-\mathrm{F})}{4\mathrm{H}(1-\mathrm{F})}, & when\ H \geq F \\ .5 - \dfrac{(\mathrm{F}-\mathrm{H})(1+\mathrm{F}-\mathrm{H})}{4\mathrm{F}(1-\mathrm{H})}, & when\ H < F \end{cases}$$

where H represents the hit-rate, and F represents the false-alarm rate.

The results are shown in Figure 4a. We found that the participants showing analytic patterns (M = .90) performed significantly better than those with holistic patterns (M = .85), t(46) = 2.24, p = .03. This result suggested that analytic patterns during face learning were beneficial for face recognition. We also computed the log-likelihoods of observing the 48 participants' learning phase eye movement data given the representative HMM of analytic patterns (Figure 3a) and examined whether they were correlated with participants' recognition performance. The log-likelihood measure reflected how similar a participant's eye movement pattern was to the representative analytic pattern. A higher value indicated higher similarity. The results (see Figure 4b), showed that although there was a positive correlation between the two measures, it did not reach significance, r(46) = .22, *p* = .13. We further verified the finding with a skipped-correlation analysis, which identified outliers and estimated the correlation after the outliers were removed (Pernet, Wilcox, & Rousselet, 2013; the analysis discovered six outliers). The result was consistent with that reported above: the correlation between the two measures was not statistically significant, r(46) = .02. The log-likelihood of observing the participants' learning phase eye movement data given the representative HMM of holistic patterns also did not correlate with their recognition performance, r(46) = .13, *p* = .37. These results suggested that although participants showing analytic patterns during face learning outperformed those showing holistic patterns in recognition, the similarities of their eye movement patterns to the representative analytic/holistic pattern were not good predictors for their recognition performance.

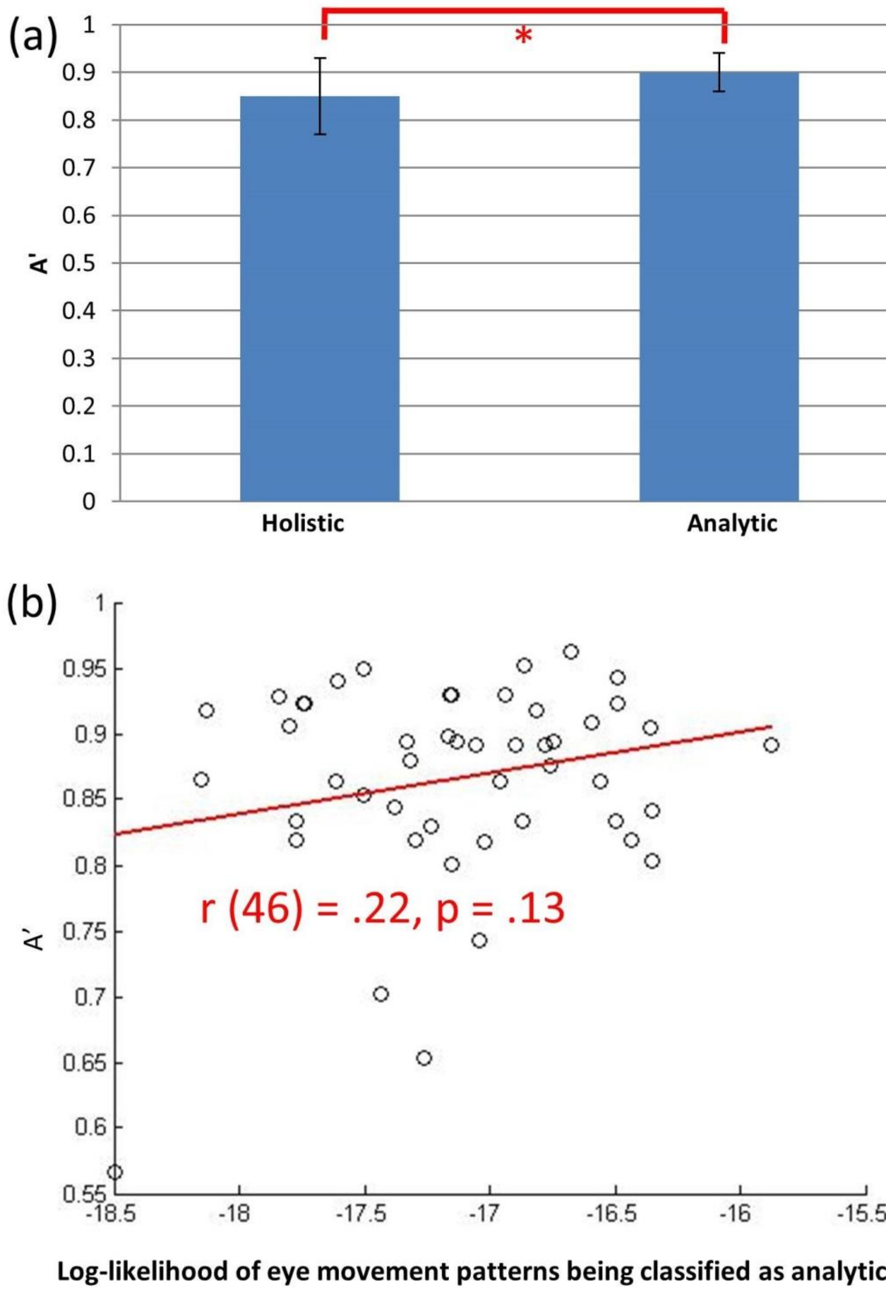


Figure 4. (a) Recognition performance of participants with different eye movement patterns during the learning phase, measured in A'. (b) The correlation between the log-likelihoods of the participants' eye movement patterns being classified as analytic and their recognition performances in A'.

Was the advantage of participants with analytic patterns in recognition performance related to the number of fixations they made during the learning phase? We found that participants with holistic patterns (M = 14.46) made a similar number of fixations to those with analytic patterns (M = 13.17), $t(46) = 1.72$, $p = .09$. This result suggested that the advantage of analytic patterns was not due to a larger number of fixations made. Instead, it may be the active sampling of information from the two eyes, the most diagnostic features for face recognition, during face learning/encoding.

Eye movement patterns during the recognition phase

To discover common eye movement patterns participants used during the recognition phase, we modeled each participant's eye movements during the recognition phase with an HMM and clustered the individual HMMs into two groups. The representative HMMs of the two groups are shown in Figure 3b.

It can be seen from Figure 3b that the four ROIDs of the holistic representative HMM were all around the center of the face. The red and blue ROIDs covered the central region of the face. The yellow ROID was at the lower part of the face, covering the tip of the nose and the mouth. The green ROID covered the nose. Participants in this group typically began a trial by looking at the center of the face with a short fixation (M = 161 ms, blue ROID). Then they looked at the center of the face with either long (M = 269 ms, red ROID) or short fixation duration (M = 161 ms, blue ROID), or occasionally (12%) they looked at the tip of the nose. Only in very rare cases (1%) would they look at the center of the nose with very long duration (M = 665 ms, green ROID). This pattern focused mainly at the center of the face, and thus we identified it as the holistic eye movement pattern.

The analytic representative HMM shown in Figure 3b had two ROIDs (the green and the red ROIDs) on the two eyes respectively. In addition, the blue ROID was at the center of the face, whereas the yellow ROID was at the tip of the nose and covered the nose and the mouth region. Participants in this group were most likely to begin a trial by looking at the center of the face with a short fixation (M = 179 ms, blue ROID). Then, they either remained looking at the center (30% of the times), or looked at the left eye (27%) or the right eye (39%) with a slightly longer fixation (left eye: M = 291 ms; right eye: M = 289 ms). They rarely (4%) looked at the nose and the mouth. When they looked at one of the eyes, their next fixation was most likely to be at the other eye, suggesting that participants in this group preferred to switch their attention between the eyes. Since this pattern showed focuses on the eyes in addition to the face center, we identified it as the analytic eye movement pattern.

The two patterns had some similarities and differences. In both patterns, participants were most likely to start a trial with a brief fixation at around the center of the face, followed by fixations with duration around 250 to 300 ms. Nevertheless, in the holistic pattern, these subsequent fixations were mostly located around the center of the face, whereas in the analytic pattern, these subsequent fixations were specifically at the two eyes. There were in total 20 participants who showed holistic patterns during the recognition phase, and 28 participants showed analytic patterns. The percentages of the participants using the two patterns did not differ significantly from each other, $\chi^2(1) = 1.33$, $p = .25$ (Table 1, recognition phase). When we compared the distribution of the participants over the two patterns during recognition with that during learning, there were significantly more participants adopting holistic patterns during learning than recognition, $\chi^2(1) = 8.30$, $p = .004$.

When we compared these patterns with those observed during the learning phase, we found that the representative holistic patterns of the two phases were significantly different. The log-likelihoods of observing the holistic eye movement data during the learning phase given the representative holistic HMM of the learning phase and the log-likelihoods of observing the same data given the representative holistic HMM of the recognition phase were significantly different, $t(33) = 4.47$, $p < .001$. The log-likelihoods of observing the holistic eye movement data during the recognition phase given the two representative holistic HMMs were also significantly different, $t(19) = 4.56$, $p < .001$. The difference between the two log-likelihoods was an approximation to the Kullback-Leibler (KL) divergence between the learning phase and recognition phase representative holistic HMMs, which was a measure of difference between two distributions (Chuk et al., 2014). Similarly, the representative analytic patterns of the two phases were significantly different. The log-likelihoods of observing the analytic eye movement data during the learning phase given the two representative analytic HMMs were significantly different, $t(13) = 2.12$, $p = .05$; similarly for those observed during the recognition phase,

$t(27) = 4.43$, $p < .001$. When we compared the number of fixations per trial that participants made during the two phases, we found that participants made significantly more fixations during the learning phase regardless of whether they used holistic patterns (M = 14.46 for learning phase, M = 6.49 for recognition phase, $t(52) = 12.55$, $p < .001$) or analytic patterns (M = 13.17 for learning phase, M = 5.85 for recognition phase, $t(40) = 10.61$, $p < .001$). We also found that the average fixation durations were significantly longer during the learning phase than the recognition phase regardless of whether participants used holistic patterns (M = 336.26 ms for learning phase, M = 246.94 ms for recognition phase, $t(40) = 4.79$, $p < .001$) or analytic patterns (M = 297.97 ms for learning phase, M = 244.27 ms for recognition phase, $t(52) = 2.91$, $p = .005$).

We also compared the durations of the fixations within the ROIDs between the two phases. The durations of the fixations at around the face center in the holistic pattern during face learning (Figure 3a, blue and red ROIDs, M = 203 ms and 311 ms respectively) were slightly longer than those observed in the holistic pattern during face recognition (Figure 3b, blue and red ROIDs, M = 161 ms and 269 ms respectively; $t(6543) = 19.37$, $p < 0.001$, and $t(7067) = 11.56$[1], $p < .001$, respectively). Similarly, the durations of the fixations at the two eyes in the analytic pattern during face learning (Figure 3a, yellow and red ROIDs, M = 327 ms and 332 ms respectively) were slightly longer than those observed in the analytic pattern during face recognition (Figure 3b, green and red ROIDs, M = 291 and 289 ms respectively; $t(4044) = 7.85$, $p < .001$, and $t(4185) = 9.09$, $p < .001$, respectively).

Regarding participants' recognition performance, we found that participants with analytic patterns (M = .89) performed significantly better than those using holistic patterns (M = .83), $t(46) = 3.13$, $p = .003$ (Figure 5a). In addition, the log-likelihoods of observing participants' recognition phase eye movements given the representative HMM of analytic patterns was positively correlated with participants' recognition performances in A', $r(46) = .35$, $p = .01$ (Figure 5b). We further verified the finding with a skipped-correlation analysis. The result was consistent with that reported above; no outlier was identified. In contrast, this correlation was not significant using the representative HMM

[1] We estimated the numbers of fixations that were responsible for the two ROIDs and used them as the sample sizes for the t-tests. The means and standard deviations are shown in the corresponding figures. The comparison of the ROIDs was done using unpaired t-tests.

of holistic patterns, r(46) = .15, p = .30. These results were consistent with our previous study (Chuk et al., 2014b; Chuk et al., submitted), suggesting that analytic eye movement patterns were beneficial for face recognition. In addition, participants using the two patterns did not differ significantly in the number of fixations made per trial, t(46) = 1.22, p = .23 (holistic patterns, M = 6.63; analytic patterns, M = 5.91) or response time (holistic patterns, M = 1.95 s; analytic patterns, M = 1.76 s), t(46) = 1.54, p = .13. This result suggested that the advantage of analytic patterns over holistic patterns was not simply because participants with analytic patterns made more fixations on the face.

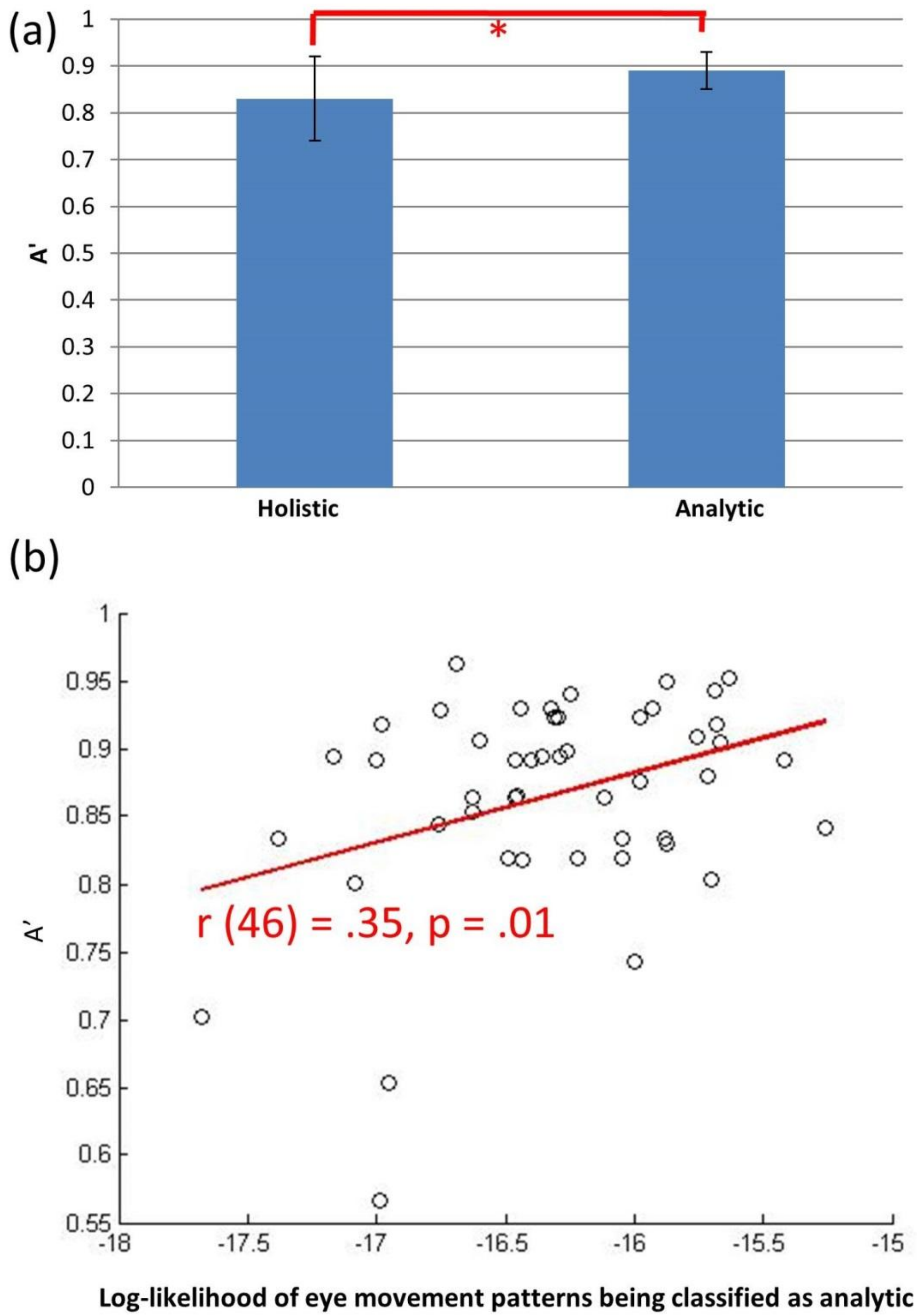
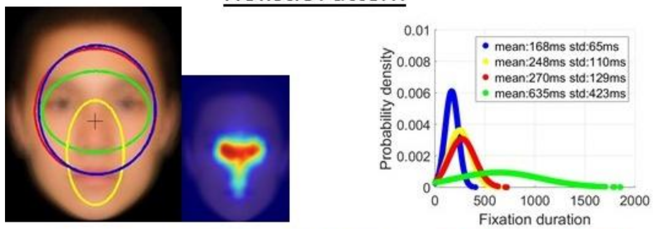


Figure 5. (a) Recognition performance of participants with different eye movement patterns during the recognition phase, measured in A'. (b) The correlation between the log-likelihoods of participants' eye movement patterns being classified as analytic and their recognition performances.

Did participants use the same eye movement patterns for the learning and recognition phases?

In the previous sections, we found that during the learning phase, a majority of participants used holistic eye movement patterns. In contrast, during the recognition phase, there were similar percentages of participant using analytic and holistic patterns. This result suggests participants might have used different eye movement patterns during the two phases. To test this, we clustered participants' learning and recognition phase HMMs into two groups to discover common patterns among them, and examined whether a majority of participants used the same patterns for face learning and recognition. The resulting patterns are shown in Figure 6. Table 2 shows the number of participants being clustered into each eye movement pattern group.

Holistic Pattern



	To Blue	To Yellow	To Red	To Green
Priors	.47	.08	.43	.02
From Blue	.30	.11	.56	.03
From Yellow	.11	.34	.52	.03
From Red	.03	.09	.79	.09
From Green	.02	.06	.58	.34

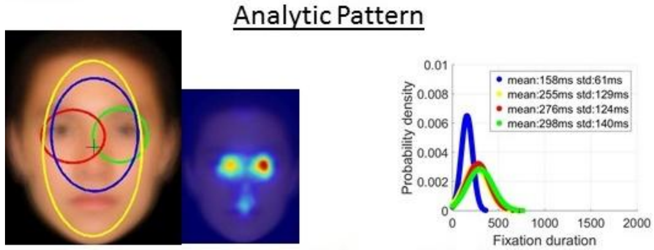
Mean location (relative to the face center) and standard deviation of ROIDs in visual angle

	Mean X	Mean Y	SD X	SD Y
Blue	.06	-.36	3.06	3.34
Yellow	.02	1.00	1.51	2.73
Red	.02	-.32	3.17	3.26
Green	.04	-.30	2.83	2.12

Mean location (relative to the face center) and standard deviation of ROIDs in face-size-normalized unit

	Mean X	Mean Y	SD X	SD Y
Blue	.01	-.06	.68	.53
Yellow	.00	.16	.33	.44
Red	.00	-.05	.69	.52
Green	.01	-.05	.63	.34

Analytic Pattern



	To Blue	To Yellow	To Red	To Green
Priors	.66	.23	.05	.06
From Blue	.22	.15	.27	.36
From Yellow	.11	.55	.14	.20
From Red	.05	.29	.12	.54
From Green	.10	.41	.37	.12

Mean location (relative to the face center) and standard deviation of ROIDs in visual angle

	Mean X	Mean Y	SD X	SD Y
Blue	.02	-.40	2.27	2.96
Yellow	-.04	.04	2.61	4.62
Red	-.66	-.34	1.66	1.40
Green	.85	-.38	1.40	1.40

Mean location (relative to the face center) and standard deviation of ROIDs in face-size-normalized unit

	Mean X	Mean Y	SD X	SD Y
Blue	.00	-.06	.50	.47
Yellow	-.01	.01	.56	.74
Red	-.14	-.05	.36	.23
Green	.19	-.06	.31	.23

Figure 6. The representative HMMs of the two common eye movement patterns discovered by clustering all participants' HMMs (including both learning and recognition phases) together. The figure shows the spatial distribution of the ROIDs and the corresponding heat map, the duration distribution of the ROIDs, and the transition probability matrix of the ROIDs. The tables below the transition matrix show the mean location (relative to the face center) and standard deviation of the ROIDs in visual angle and in face-size-normalized unit.

Learning phase			
(a)	Female	Male	**Total**
Holistic pattern	26	9	**35**
Analytic pattern	9	4	**13**
(b)	Caucasian	Asian	
Holistic pattern	18	17	**35**
Analytic pattern	6	7	**13**
Recognition phase			
(a)	Female	Male	**Total**
Holistic pattern	16	4	**20**
Analytic pattern	19	9	**28**
(b)	Caucasian	Asian	
Holistic pattern	12	8	**20**
Analytic pattern	12	16	**28**

Table 2. The number of participants being clustered into each eye movement pattern group (analytic vs. holistic) with a breakdown by gender (a) and by race (b), using the representative HMMs in Figure 6.

It can be seen that in the holistic representative HMM in Figure 6, the red, blue, and green ROIDs centered at the bridge of the nose, and the yellow ROID covered the nose and the mouth region. In this pattern, participants were most likely to begin a trial with a longer (M = 270 ms) or a shorter (M = 168 ms) fixation at the center of the face. Afterwards, they typically made a long (M = 270 ms) fixation at around the center of the face. Occasionally, they looked at the center of the face with much longer durations (M = 635 ms, green ROID), or the tip of the nose/mouth region (yellow ROID). This pattern reflected a focus at the center of the face, and thus we identified it as the holistic eye movement pattern.

The analytic representative HMM shown in Figure 6 reflected a different eye movement pattern. The blue and the yellow ROIDs both centered at the bridge of the nose, whereas the red and the green ROIDs were located at the left and the right eye respectively. In this pattern, participants were most likely to begin a trial with a short fixation at the center of the face (M = 158 ms, blue ROID), followed by a longer fixation on either the right eye (M = 298 ms) or the left eye (M = 276 ms). Sometimes they remained looking at the center with either a short (M = 158 ms, blue ROID) or a longer (M = 255 ms, yellow ROID) fixation. Since this pattern showed specific focuses on the two eyes in addition to the face center, we identified it as the analytic eye movement pattern.

We found that 35 participants' learning phase HMMs were clustered into the holistic pattern and 13 were clustered into the analytic pattern. For the recognition phase HMMs, 20 participants' HMMs were clustered into the holistic pattern and 28 were clustered into the analytic pattern (Table 2). As summarized in Table 3 below, 19 (about 40%) participants used different eye movement patterns between the two phases, and 29 participants used the same patterns between the two phases. The percentages of participants using the same or different patterns between the two phases did not differ significantly, $\chi^2(1) = 2.08$, $p = .15$. Interestingly, among participants who used different patterns during the two phases, 90% of them (17/19) switched their patterns from holistic at learning to analytic at recognition.

Pattern switch		*recognition phase*		
		same	different	**Total**
learning	holistic	18	17	**35**
phase	analytic	11	2	**13**
Total		**29**	**19**	**48**

Table 3. Number of participants switched patterns during the two phases.

To test whether participants' perceptuomotor memory during face learning played an important role in their recognition performance, as suggested by the scan path theory, we examined whether participants who used the same eye movement patterns between face learning and recognition outperformed those who used different patterns in face recognition. The results showed that the two groups did not differ significantly in recognition performance (participants who used different patterns, $M = .87$; participants who used same patterns, $M = .86$), $t(46) = .36$, $p = .72$. In a separate analysis, we performed a 2 x 2 ANOVA with learning phase eye movement pattern (holistic vs. analytic) and recognition phase eye movement pattern (holistic vs. analytic) as independent variables and recognition performance in A' as the dependent variable. We found that the two factors did not interact with each other, $F(1, 44) = .06$, $p = .82$, suggesting that whether participants changed their eye movement patterns between the learning and the recognition phases did not significantly modulate recognition performance. Note however that this analysis was based on unequal numbers of participants in each condition, as shown in Table 3. We also examined whether participants' recognition performance was correlated with the similarity between their learning phase and recognition phase eye movement patterns. To do this, for each participant, we calculated the log-likelihoods of observing the participant's recognition phase eye movement data given his/her learning phase and recognition phase HMMs. The difference between the two log-likelihoods represented the KL-divergence between the learning phase and recognition phase HMMs, a measure of similarity between the two eye movement patterns. We found that this similarity measure did not correlate with recognition performance, $r(46) = .18$, $p = .22$. Similarly, the correlation using the participants' learning phase eye movement data was not significant, $r(46) = .11$, $p = .44$. These results suggested that the similarity between learning phase and recognition phase eye movement patterns did not predict recognition performance.

Discussion

In this study, we aimed to examine the relationship between eye movement patterns during face learning and recognition, and its association with recognition performance in a face recognition memory task. To reflect individual differences in both spatial and temporal dimensions of eye movements in our data analysis, we used a hidden Markov model (HMM) based approach (Chuk et al., 2014), in which each participant's eye movement pattern was modeled with an HMM. The hidden states of the HMMs represented regions of interest and duration (i.e., ROID) of participants' fixations. The eye movements among these ROIDs were summarized with a transition matrix in the model. This information was estimated from participants' eye movement data in a completely data-driven fashion. Individual HMMs then could be clustered according to their similarities to discover common patterns shared by individuals. The similarity between an individual's eye movement pattern to a common pattern discovered through clustering could be calculated as the likelihood of the individual pattern being classified as the common pattern. This similarity measure then could be used to examine the association between eye movement patterns and recognition performance. Note that in contrast to the HMMs used in our previous studies (e.g., Chuk et al., 2014), the HMM used in the current study was improved in two aspects. First, the number of hidden states was determined through model selection instead of pre-specified. Second, we included fixation duration information in addition to fixation location information. This is to reflect the previous finding that eye movements during face learning and recognition differed in fixation duration (Hsiao & Cottrell, 2008). The new model thus was able to more accurately summarize a participant's eye movement behavior in a cognitive task.

Our results showed that both holistic (i.e., looking mainly at the face center) and analytic eye movement patterns (i.e., looking specifically at the two eyes in addition to the face center) could be observed during face learning and recognition. Nevertheless, the holistic and analytic patterns observed during face learning differed significantly from those observed during face recognition. Eye movements during the learning phase occasionally involved long fixations at around the center of the face, which was rarely the case during the recognition phase. In addition, the fixations during learning were in general longer and more numerous than those observed during recognition. Interestingly, we found that significantly more participants adopted holistic patterns during face learning than recognition. Combined, these results suggested that in general participants showed different eye movement patterns between face learning and recognition, demonstrating different cognitive processes involved for information encoding and retrieval.

Hsiao and Cottrell (2008) observed that when comparing the first three fixations during the learning and recognition phases, during learning participants' fixation duration gradually increased from the first to the third fixations, whereas during the recognition phase, the first three fixations were of similar durations. This pattern was in general consistent with our results. During face learning, most participants adopted holistic patterns. In the representative holistic pattern during learning (Figure 3a), participants typically started a trial with a short fixation at the face center (M = 203 ms, blue ROID), and gradually transited to longer fixations at the face center at third fixation (M = 311 ms, red ROID). Whereas during face recognition, in both holistic and analytic patterns, participants typically started with a short fixation (M ~ 170 ms), followed by slightly longer fixations (M ~ 278 ms) at both the second and third fixations. In contrast to our finding, Blais et al. (2008) and Caldara et al. (2010) found that participants' fixation durations did not differ between the learning and recognition phases using group-level analysis. We speculate that this discrepancy may be due to substantial individual differences in eye movement pattern during the two phases. Our approach allowed us to discover different patterns within each phase, and compare corresponding ROIDs in similar patterns across the two phases. And thus we were able to better discover this difference in fixation duration between the two phases.

Although the holistic and analytic eye movement patterns during the learning and recognition phases differed in fixation duration, we found that in both phases, participants with analytic patterns outperformed those with holistic patterns in recognition performance. This finding was consistent with Sekiguchi's (2011) finding that participants who performed better in face recognition moved their eyes between the left and right eyes more often during face learning as compared with those who performed worse. Note that this advantage of analytic patterns was not because participants using analytic patterns made more fixations per trial than those using holistic patterns, as the two groups of participants did not differ significantly in number of fixations made per trial either for face learning or recognition. Instead, this advantage of analytic patterns was likely to be due to active information encoding and retrieval from the two eyes, suggesting that information about the two eyes is important for face recognition. Consistent with this finding, the two eyes have been reported to be the most diagnostic features participants used for face recognition (e.g., Gosselin & Schyns, 2001; Vinette et al., 2004). The eyes also have been proposed to provide important signals for the direction of social attention (e.g., Langton, Watt, & Bruce, 2000). In addition, as compared with whole faces, eyes presented in isolation are shown to elicit larger N170 ERP amplitude, an electrophysiological marker proposed to reflect the neural mechanism for face detection, suggesting the importance of eyes in face perception (Bentin, Allison, Puce, Perez, & McCarthy, 1996; see also Taylor, Itier, Allison, & Edmonds, 2001; Taylor, Edmonds, McCarthy, & Allison, 2001). Although analytic eye movement patterns during both face learning and recognition seemed to be beneficial for face recognition, we found that participants' recognition performance was positively correlated with the log-likelihood of participants' eye movements being classified as analytic during the recognition phase, but not with that during the learning phase. This finding suggested that eye movement patterns during the recognition phase may be a better predictor for participants' recognition performance than those during the learning phase.

Miellet, Caldara, and Schyns (2011) showed that during face recognition, participants' eye fixations on the eyes of the face images were associated with perception of local information, whereas those at the center of the face were associated with perception of global information. According to this finding, participants using the analytic and holistic eye movement patterns identified in the current study may engage different types of information processing in face recognition. More specifically, participants with holistic patterns (i.e., looking mainly at the face center) may have primarily engaged in global/configural face processing, whereas those with analytic patterns (i.e., focusing on the individual eyes in addition to the face center) may have engaged in both global and local/featural face processing. While global/configural information was reported to play an important role in face recognition (e.g. Bartlett & Searcy, 1993; Leder & Bruce, 1998), most recent studies have suggested that both local/featural and global/configural information are important for recognizing faces (e.g., Burton, Schweinberger, Jenkins, & Kaufmann, 2015; Cabeza & Kato, 2000; Sandford & Burton, 2014). Consistent with this finding, in automatic face recognition in computer vision, the best performing algorithms made use of both local and global representations of the faces (Bonnen, Klare, & Jain, 2013; Ding, Shu, Fang, & Ding, 2010). Together, these findings suggested that active retrieval of both global and local face representations through analytic eye movement patterns may be optimal for face recognition.

In order to examine whether individual participants used the same or different eye movement patterns between the learning and the recognition phases, in a separate analysis we clustered participants' learning and recognition phase HMMs together into two groups to discover common patterns shared between the two phases. The resulting two representative HMMs (Figure 6) showed similar characteristics as the holistic and analytic patterns discovered when we clustered participants' patterns in the learning and recognition phases separately. We then examined whether individual participants used the same or different patterns between the learning and recognition phases. We found that about 40% of the participants used different eye movement patterns between the learning and recognition phases, and the percentages of the participants using the same or different patterns did not differ significantly from each other. Interestingly, among those who used different patterns between learning and recognition, 90% of them switched from holistic at learning to analytic at recognition, suggesting that analytic patterns were preferred during recognition. These results showed that participants do not necessarily use the same eye movement patterns for face learning and recognition, This finding was in contrast to previous studies that observed similar eye movement patterns between the learning and recognition phases using group-level eye movement data analysis (e.g., Blais et al., 2008; Caldara et al., 2010). This individual difference in the similarity of eye movements between learning and recognition may have been obscured in the group-level data analysis. This phenomenon demonstrated well the advantage of our approach for data analysis at the individual level.

According to the scan path theory (Noton & Stark, 1971a; 1971b), recapitulation of the eye movement/perceptuomotor pattern produced during learning is necessary for recognition to be successful. If perceptuomotor memory elicited by eye movements does play an important role for recognition performance, we would expect that participants who used the same eye movement pattern between the learning and recognition phases outperformed those who used different eye movement patterns in the recognition task. Nevertheless, the results of our analysis did not support this hypothesis. We found that participants who showed the same or different eye movement patterns between the two phases did not differ significantly in their recognition performance. In addition, the similarity between their eye movement patterns during the learning and the recognition phases did not significantly correlate with their recognition performance. Instead, we found that analytic eye movement patterns during the recognition phase, which focused on the two eyes in addition to the face center, seemed to be the best predictor for participants' recognition performance. This phenomenon suggested that retrieval of the most diagnostic features for recognition is more important than recapitulation of the perceptuomotor cycles/eye movements produced during learning in visual recognition. To confirm the speculation that eye fixations at more diagnostic features for recognition lead to better recognition performance, future work will directly manipulate participants' eye movement patterns (such as through cueing or training paradigms; e.g., Hills & Lewis, 2011) and examine whether it modulates their recognition performance.

Note that in the current study, we used the same images for old faces during the learning and recognition phases, following a majority of face recognition studies in the literature (e.g., Barton et al., 2006; Hayward, Rhodes, & Schwaninger, 2008; Henderson, Williams, & Falk, 2005; Hsiao & Cottrell, 2008). However, real-life face recognition typically involves recognizing faces under different conditions, such as different orientations, expressions, or lighting conditions. Future work will examine how these different task demands modulate the association between participants' eye movement patterns and performance in face recognition.

In summary, through analyzing eye movement data at the individual level using the HMM based approach, here we showed that both holistic and analytic eye movement patterns could be observed during face learning and recognition. Eye movements during learning generally involved longer fixation duration than those during recognition. During both face learning and recognition, participants who showed analytic patterns performed better than those with holistic patterns in the recognition task, although a significant correlation between eye movement patterns and recognition performance was only observed for eye movements during the recognition phase. This finding suggested that the retrieval of diagnostic features for recognition, such as the eyes, is a good predictor for performance in face recognition. In contrast to the scan path theory, which posits eye movements produced during learning have to be repeated during recognition for the recognition to be successful, we found that participants used the same eye movement pattern for face learning and recognition did not differ from those used different patterns in recognition performance. In addition, the similarity between the eye movement patterns during face learning and recognition did not correlate with recognition performance. These results suggested that perceptuomotor memory elicited by eye movement patterns during learning does not play an important role in recognition. In contrast, it is the retrieval of diagnostic information during recognition that is essential for recognition to be successful. This finding has very important implications for ways to improve recognition performance in both healthy and clinical populations.

Acknowledgements

Janet H. Hsiao and Antoni B. Chan are grateful to the Research Grant Council of Hong Kong (Project number 17402814 to J.H. Hsiao and CityU 110513 to A.B. Chan) and HKU Seed Funding Programme for Basic Research (Project number 201311159131 to J.H. Hsiao). We are grateful to the editor and two anonymous reviewers for their helpful comments.

References

Andrews, T. J., & Coppola, D. M. (1999).Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision research*, *39*(17), 2947-2953.

Antrobus, J. S., Antrobus, J. S., & Singer, J. L. (1964).Eye movements accompanying daydreaming, visual imagery, and thought suppression. *The Journal of Abnormal and Social Psychology*, *69*(3), 244.

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cogn. Psychol.*, *25*(3), 281-316.

Barton, J. J., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. Perception, 35(8), 1089-1105.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of cognitive neuroscience*, *8*(6), 551-565.

Bishop, C. M. (2006).Pattern Recognition. *Machine Learning*.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS One*, *3*(8), e3022.

Bonnen, K., Klare, B., & Jain, A. (2013). Component-based representation in automated face recognition. IEEE T. Inf. Foren. Sec., 8(1), 239-253.

Burton, A. M., Schweinberger, S. R., Jenkins, R., & Kaufmann, J. M. (2015). Arguments against a configural processing account of familiar face recognition. Perspectives on Psychological Science, 10(4), 482-496.

Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural and configural processing to face recognition. Psychol. Sci.,11(5), 429-433.

Caldara, R., Zhou, X., &Miellet, S. (2010).Putting culture under the ‘spotlight’reveals universal information use for face recognition. *PLoS One*,*5*(3), e9708.

Castelhano, M. S., & Henderson, J. M. (2008).Stable individual differences across images in human saccadic eye movements. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*,*62*(1), 1.

Charness, N., Reingold, E. M., Pomplun, M., &Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & cognition*, *29*(8), 1146-1152.

Cheng Z., Chuk, T., Hayward, W. G., Chan, A. B., & Hsiao, J. H. (2015). Global and local priming evoke different face processing strategies: evidence from an eye movement study. Journal of Vision, 15(12), 154.

Chuk, T., Chan, A. B., & Hsiao, J. H. (2014a).Understanding eye movements in face recognition using hidden Markov models. *Journal of vision*, *14*(11), 8.

Chuk, T., Luo, A. X., Crookes, K., Hayward, W. G., Chan, A. B., & Hsiao, J. (2014b). Caucasian and Asian eye movement patterns in face recognition: A computational exploration using hidden Markov models. *Journal of vision*,*14*(10), 1212-1212.

Chuk, T., Crookes, K., Hayward, W. G., Chan, A. B., & Hsiao, J. H. (submitted). Hidden Markov model analysis reveals the association between eye movement patterns and face recognition performance across cultures.

Coviello, E., Chan, A. B., &Lanckriet, G. R. (2014).Clustering hidden Markov models with variational HEM. *The Journal of Machine Learning Research*, *15*(1), 697-

747.

Ding, L., Shu, C., Fang, C., & Ding, X. (2010). Computers do better than experts matching faces in a large population. In Cognitive Informatics (ICCI), 2010 9th IEEE International Conference, 280-284.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. Neuropsychologia, 44(4), 576-585.

Glen, F. C., Crabb, D. P., Smith, N. D., Burton, R., &Garway-Heath, D. F. (2012). Are Certain Eye Movement Patterns Linked To Better Face Recognition Performance In Patients With Central Glaucomatous Visual Field Loss?. *Investigative Ophthalmology & Visual Science*, *53*(14), 5623-5623.

Goldinger, S. D., He, Y., &Papesh, M. H. (2009). Deficits in cross-race face learning: Insights from eye movements and pupillometry. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1105.

Gosselin, F., &Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision research*, *41*(17), 2261-2271.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving guiding attention guides thought. *Psychological Science*, *14*(5), 462-466.

Hayward, W. G., Rhodes, G., & Schwaninger, A. (2008). An own-race advantage for components as well as configurations in face recognition. Cognition, 106(2), 1017-1027.

Henderson, J. M. (2003). Human gaze control during real-world sceneperception. *Trends in Cognitive Sciences*, *7*, 498-504.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & cognition*, *33*(1), 98-106.

Heremans, E., Helsen, W. F., &Feys, P. (2008). The eyes as a mirror of our thoughts: quantification of motor imagery of goal-directed movements through eye movement registration. *Behavioural brain research*, *187*(2), 351-360.

Hills, P. J. & Lewis, M. B. (2011). Reducing the own-race bias in face recognition attentional shift using fixation crosses preceding the lower half of a face. *Vis. Cogn.*, *19*, 313- 339.

Hsiao, J. H. W., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science*, *19*(10), 998-1006.

Hyönä, J., Lorch Jr, R. F., &Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns.*Journal of Educational Psychology*, *94*(1), 44.

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.

Kelly, D. J., Jack, R. E., Miellet, S., De Luca, E., Foreman, K., & Caldara, R. (2011). Social experience does not abolish cultural diversity in eye movements. *Frontiers in psychology*, *2*, 95.

Laeng, B., & Teodorescu, D. S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*,*26*(2), 207-231.

Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in cognitive sciences*, *4*(2), 50-59.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Q. J. Exp. Psychol. A, 51*(3), 449–473.

Martin, V., Hurn, S., & Harris, D (2012). *Econometric Modelling with Time Series: Specification, Estimation and Testing*. Cambridge University Press.

Mehoudar, E., Arizpe, J., Baker, C. I., &Yovel, G. (2014).Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of vision*, *14*(7), 6-6.

Miellet, S., Caldara, R., & Schyns, P. G. (2011). Local Jekyll and Global Hyde The Dual Identity of Face Identification. Psychological science, 22(12), 1518-1526.

Noton, D., & Stark, L. (1971a). Scan paths in eye movements during pattern perception. *Science*, *171*, 308-311.

Noton, D., & Stark, L. (1971b). Scan paths in saccadic eye movements swhile viewing and recognizing patterns. *Vision Research*, *11*,929-942.

Oates, T., Firoiu, L., & Cohen, P. R. (2001). Using Dynamic Time Warping to Bootstrap HMM-Based Clustering of Time Series. *Sequence Learning: Paradigms, Algorithms, and Applications*. pp. 35-52, Springer-Verlag.

Ohl, S., Brandt, S. A., &Kliegl, R. (2013).The generation of secondary saccades without postsaccadic visual feedback. *Journal of vision*, *13*(5), 11-11.

Pernet, C. R., Wilcox, R. R., & Rousselet, G. A. (2013). Robust correlation analyses: false positive and power validation using a new open source Matlab toolbox. *Frontiers in psychology*, *3*, 606.

Peterson, M. F., & Eckstein, M. P. (2013). Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychological science*, *24*(7), 1216-1225.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372- 422

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*(1), 86-90.

Sandford, A., & Burton, A. M. (2014). Tolerance for distorted faces: Challenges to a configural processing account of familiar face recognition. Cognition, 132(3), 262-268.

Sekiguchi, T. (2011). Individual differences in face memory and eye fixation patterns during face learning. *Actapsychologica*, *137*(1), 1-9.

Seo, T.-K., Kishino, H., & Thorne, J. L. (2005). Incorporating gene-specific variation when inferring and evaluating optimal evolutionary tree topologies from multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America*, 102(12), 4436–4441.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011).Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, *27*(2), 251-272.

Taylor, M. J., Edmonds, G. E., McCarthy, G., & Allison, T. (2001). Eyes first! Eye processing develops before face processing in children. *Neuroreport*,*12*(8), 1671-1676.

Taylor, M. J., Itier, R. J., Allison, T., & Edmonds, G. E. (2001). Direction of gaze effects on early face processing: eyes-only versus full faces. *Cognitive Brain Research*, *10*(3), 333-340.

Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520-522.

Tovée, M. J. (1996). *An introduction to the visual system*.Cambridge University Press.

Vinette, C., Gosselin, F., &Schyns, P. G. (2004).Spatio-temporal dynamics of face recognition in a flash: It’s in the eyes. *Cognitive Science*, *28*(2), 289-301.

Waters, A. J., & Underwood, G. (1998). Eye movements in a simple music reading task: A study of expert and novice musicians. *Psychology of Music*,*26*(1), 46-60.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., ... & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. Proceedings of the National Academy of sciences, 107(11), 5238-5241.

Yarbus, A. L. (1967). *Eye movements during perception of complex objects*(pp. 171-211). Springer US.