Cognition 211 (2021) 104616

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Do portrait artists have enhanced face processing abilities? Evidence from hidden Markov modeling of eye movements

Janet H. Hsiao [a,b,*], Jeehye An [a], Yueyuan Zheng [a], Antoni B. Chan [c]

[a] Department of Psychology, University of Hong Kong, Hong Kong Special Administrative Region
[b] The State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong Special Administrative Region
[c] Department of Computer Science, City University of Hong Kong, Hong Kong Special Administrative Region

A R T I C L E I N F O

Keywords:
Face-drawing experience
Face perception
Eye movement
Hidden Markov model
EMHMM

A B S T R A C T

Recent research has suggested the importance of part-based information in face recognition in addition to global, whole-face information. Nevertheless, face drawing experience was reported to enhance selective attention to the eyes but did not improve face recognition performance, leading to speculations about limited plasticity in adult face recognition. Here we examined the mechanism underlying the limited advantage of face drawing experience in face recognition through the Eye Movement analysis with Hidden Markov Models (EMHMM) approach. We found that portrait artists showed more eyes-focused eye movement patterns and outperformed novices in face matching, and participants' drawing rating was correlated with both eye movement pattern and performance. In contrast, portrait artists did not outperform novices and did not differ from novices in eye movement pattern in either the face recognition or part-whole tasks, although the eyes-focused pattern was associated with better recognition performance and longer response times in the whole condition relative to the part condition. Interestingly, in contrast to the face recognition and part-whole tasks, participants' performance in face matching was predicted by their drawing rating but not eye movement pattern. These results suggested that artists' advantage in face processing is specific to tasks similar to their drawing experience such as face matching, and may be related to their better ability in extracting identity-invariant information between two faces rather than more eyes-focused eye movement patterns.

1. Introduction

The ability to efficiently and accurately recognize or identify a face is an essential skill in daily life. Due to its importance, how humans recognize faces has been extensively studied. Some classic effects in face perception have suggested that humans process faces as a whole instead of based on individual facial parts. For example, the part-whole effect shows that when identifying a facial part of a learned face (such as the eyes or the nose), participants perform better when it is presented in a whole face context than presented alone (e.g., Farah, Wilson, Drain, & Tanaka, 1998; Tanaka & Farah, 1993). The composite face effect refers to the phenomenon that two identical top half-faces are perceived as different when they are paired with different bottom half-faces, demonstrating that the perception of facial parts is dependent on the whole face context as the result of holistic face processing (e.g., Hole, 1994; Young, Hellawell, & Hay, 1987). Also, the face inversion effect demonstrates our enhanced sensitivity to configuration of upright faces, in particular the spatial relations among facial parts, as compared with inverted faces (e.g., Kemp, McManus, & Pigott, 1990; Bartlett & Searcy, 1993; Searcy & Bartlett, 1996; Freire, Lee, & Symons, 2000; Barton, Keenan, & Bass, 2001). Although these classic effects may involve different perceptual mechanisms (e.g., Rezlescu, Susilo, Wilmer, & Caramazza, 2017), they consistently suggest that global (whole-face) configural information plays an important role in face recognition. Consistent with these findings, the ability to draw long-range spatial relations among facial parts is important for face drawing accuracy (Ostrofsky, Cohen, & Kozbelt, 2014), and face inversion and misalignment of top and bottom half-faces are shown to particularly impair this ability (Ostrofsky, Kozbelt, Cohen, Conklin, & Thomson, 2016; Ostrofsky, Pletcher, & Smith, 2020). More recently, some researchers reported that local (relative to the face size) featural information also contributes significantly, suggesting that both featural and configural information are important for face recognition (e.g., Burton, Schweinberger, Jenkins, & Kaufmann, 2015; Cabeza & Kato, 2000; Harris & Aguirre, 2008;

* Corresponding author at: Department of Psychology, University of Hong Kong, Pokfulam Road, Hong Kong.
E-mail address: jhsiao@hku.hk (J.H. Hsiao).

https://doi.org/10.1016/j.cognition.2021.104616
Received 9 July 2020; Received in revised form 24 January 2021; Accepted 27 January 2021
Available online 13 February 2021

Harris & Nakayama, 2008; Hayward, Crookes, & Rhodes, 2013; Lobmaier & Mast, 2007; Mondloch et al., 2010; O'Toole, Deffenbacher, Valentin, & Abdi, 1994). In particular, forensic facial identification examiners, who outperformed untrained participants in challenging face identity matching tasks, showed a reduced face inversion effect, suggesting that the ability to process local featural information in addition to global configural information may be essential for achieving optimal face identification performance (White, Phillips, Hahn, Hill, & O'Toole, 2015). Together these findings suggest that optimal face recognition performance involves a combination of global and local face processing.

Consistent with human subject studies, in computer vision research, state-of-the-art automatic face recognition solutions that outperform humans typically involve a combination of local (component-based/features-based processing) and global representations (relative to the stimulus size. E.g. Ding, Shu, Fang, & Ding, 2010; Bonnen, Klare, & Jain, 2013). In addition to faces, the advantage of combining global and local methods has been reported in automatic recognition of fingerprints (Jain, Chen, & Demirkus, 2007), finger–knuckle-prints (Zhang, Zhang, Zhang, & Zhu, 2011), palm prints (Zhang, Zuo, & Yue, 2012), handwritten Chinese characters (Gao, Uozumi, & Chen, 2005), Korean characters (Lee & Kim, 2008), and Tibetan characters (Ma & Wu, 2012). These findings suggest that both global and local information processing are required for achieving optimal performance in the recognition of faces and visual objects.

1.1. Eye movement patterns in face recognition

The importance of local featural information in addition to global configural information has also been demonstrated in eye fixation behavior during face recognition. For example, people often look towards the eye region in social scenes (e.g., Birmingham, Bischof, & Kingstone, 2008a, 2008b), and looking at the eyes is shown to be associated with local face processing, or more specifically, the use of high spatial frequency information of the eye region (Miellet, Caldara, & Schyns, 2011). The eye region has been reported to be the most diagnostic feature for face recognition (Gosselin & Schyns, 2001). Looking towards the eye region (Davis et al., 2017), or more precisely, just below the eyes (Peterson & Eckstein, 2012), has been shown to predict better face recognition performance (note however that Mehoudar, Arizpe, Baker, & Yovel, 2014, reported that individual eye movement patterns were not predictive of performance in face recognition). In particular, through the Eye Movement analysis with Hidden Markov Modeling (EMHMM) approach, Chuk, Chan, and Hsiao (2014) discovered two common eye movement patterns (i.e., eye gaze locations and transitions) in face recognition through clustering individual eye movement patterns according to their similarities: a nose-focused pattern where most of the eye fixations are located around the nose/face center, and an eyes-nose pattern where most of the eye gazes are directed to the eye region in addition to the face center. Participants adopting the eyes-nose pattern were found to have better face recognition performance than those using the nose-focused pattern (Chuk, Chan, & Hsiao, 2017; Chuk, Crookes, Hayward, Chan, & Hsiao, 2017), whereas the nose-focused pattern is associated with cognitive decline in older adults, particularly in visual attention ability and executive function (Chan, Chan, Lee, & Hsiao, 2018; see also Zhang, Chan, Lau, & Hsiao, 2019). Note that eye movement pattern alone does not tell us whether participants are engaging global or local attention (relative to the stimulus size). Thus, to examine the associations between eye fixation behavior and information use during face recognition, Miellet et al. (2011) manipulated spatial frequency content of a face image so that participants would see different face identities from the same face image when they focused on high spatial frequency information/local details and low spatial frequency/global information of the face. Through this design, they showed that when participants look at the eyes, they engage in high spatial frequency/local face processing, whereas when they look at the nose, they engage in low spatial frequency/global face processing. Consistent with this finding, asking participants to match small letters in hierarchical patterns (Navon, 1977) as local attention priming is reported to increase their likelihood of using the eyes-nose eye movement pattern (as opposed to the nose-focused pattern), suggesting that looking at the eyes during face recognition is associated with engagement of local attention (Cheng, Chuk, Hayward, Chan, & Hsiao, 2015). Accordingly, the eyes-nose pattern, where eye gazes are directed to both the eyes and the nose, may involve both local and global facial information processing, and consequently leads to better face recognition performance (Chuk et al., 2017). Thus, engagement of local featural processing in addition to global configural processing may be optimal for face recognition.

1.2. Face drawing expertise and face processing

Recent research has suggested that drawing practice involves a fine-grained procedure with a mixture of global and local information processing, which may enhance the ability to identify and integrate local components to support the understanding of global structure (Perdreau and Cavanagh, 2013) and flexibility in switching between global and local selective attention (Chamberlain & Wagemans, 2015; Kozbelt & Ostrofsky, 2018). Indeed, drawing expertise is shown to be associated with better performance on tasks requiring local visual processing without a reduction in global visual processing ability (Chamberlain, McManus, Riley, Rankin, & Brunswick, 2013). Thus, expertise in drawing may enhance face recognition performance. Consistent with this speculation, recent research on perceptual expertise acquisition has suggested that motor learning experience may enhance selective attention to local features (relative to the stimulus size) and in turn enhance recognition performance. For example, using the composite paradigm, Tso, Au, and Hsiao (2014) showed that as compared with non-Chinese readers, proficient Chinese readers who have limited Chinese character writing experience (Limited Writers) showed increased holistic processing in Chinese character perception whereas expert Chinese readers who are proficient in both reading and writing Chinese characters (Writers) showed reduced holistic processing. In addition, Writers had significantly shorter character naming time than Limited writers, suggesting more efficient character processing due to writing experience (see also Tso, Au, & Hsiao, 2012). Consistent with this finding, Chinese children's character copying and dictation performance is reported to be correlated with reading performance (McBride-Chang, Chung, & Tong, 2011; Tan, Spinks, Eden, Perfetti, & Siok, 2005), and writing practice is shown to strengthen Chinese character recognition (Guan et al., 2015). Brain imaging data also suggest that writing plays an important role in shaping the neural representation specialized for reading (e.g., Longcamp, Anton, Roth, & Velay, 2003; Siok, Perfetti, Jin, & Tan, 2004). These findings suggest that sensorimotor experience enhances selective attention to local parts, which in turn facilitates recognition.

Although faces are processed holistically in general, face drawing expertise may enhance the ability to attend to local facial features in addition to global configural processing, which may be beneficial to face processing performance. Indeed, a similar modulation effect of sensorimotor experience on holistic processing has been reported in face perception: Zhou, Cheng, and Wong (2012) reported that art students with face drawing experience had a reduced composite face effect, indicating reduced holistic face processing, as compared with non-drawers. This result suggests that drawing experience may enhance processing of local facial features. Consistent with this finding, eye tracking studies of portrait drawing artists have shown that artists capture visual information detail by detail instead of in a holistic manner with stable oculomotor fixations of long duration (Miall & Tachalenko, 2001; Tchalenko, Dempere-Marco, Hu, & Yang, 2003). Nevertheless, art students and non-drawers did not differ in face recognition performance regardless of their difference in holistic face processing. This finding is in contrast to the observations from forensic facial identification examiners, who demonstrated a reduced face inversion effect and enhanced face identity matching performance (White et al., 2015). It is also in contrast to the findings from research on Chinese character recognition and drawing expertise, where writers showed reduced holistic processing and more efficient character recognition as compared with limited writers (Tso et al., 2014), and expert drawers showed both enhanced local visual processing and memory of objects they drew (Perdreau & Cavanach, 2015). Indeed, although portrait artists have often been reported to have better perceptual discrimination abilities for faces, evidence supporting their advantage in face recognition performance has been very limited (e.g., Devue & Barsics, 2016; Dolzycka, Herzmann, Sommer, & Wilhelm, 2014; Tree, Horry, Riley, & Wilmer, 2017). Tree et al. (2017) have speculated that due to our abundant experience in face recognition, adult face recognition performance may have reached a capacity limit determined mostly by genes, leaving little plasticity for further improvement through training. Consistent with this speculation, recent twin studies have shown that monozygotic twins had a much larger correlation of scores in face recognition than dizygotic twins, suggesting a significant genetic contribution in face recognition ability (Shakeshaft & Plomin, 2015; Wilmer et al., 2010; Zhu et al., 2010).

In contrast to studies with portrait artists, experienced forensic facial identification examiners were reported to have superior face identification matching performance (White et al., 2015). Note however that White et al. (2015) used a simultaneous face identity matching task, which matched well with what forensic facial identification examiners typically performed at work. Thus, forensic facial identification examiners' advantage in face processing may be limited to tasks similar to what they were trained for such as face identity matching tasks. Similarly, when drawing a portrait, portrait artists look back and forth between the model and the drawing to acquire important global and local visual information from the model and match it with the lines that already exist on the paper to continue the drawing (Tchalenko et al., 2003). Thus, their portrait drawing experience may involve discerning and matching identity-invariant, global and local facial information between two faces, which is also required in a face matching task. In contrast, face recognition involves remembering and retrieving idiosyncratic and diagnostic facial information of a face to distinguish it from all other face identities in the memory, which may differ significantly from portrait artists' typical experience with faces. Thus, portrait artists' better abilities in selectively attending to facial parts and in face processing may be limited to tasks similar to their drawing experience such as face identity matching, but not in face recognition. Consequently, they may have better performance and show a more eyes-focused eye movement pattern than novices only in face identity matching, but not in face recognition memory tasks. In other words, drawing experts' advantage in face processing may be task-specific rather than task-general.

Here we recruited portrait artists and age-matched novices and examined their eye movement patterns and performance in a simultaneous face identity matching and a face recognition memory task. To further examine portrait artists' ability in the recognition of face parts as opposed to whole faces as compared with novices, we also included a part-whole task. We expected that, as compared with novices, portrait artists may show better performance and better abilities in attending to facial parts, as reflected in a more eyes-focused eye movement pattern, only during simultaneous face identity matching, a task that was similar to their drawing experience, but not in the other two face tasks.

2. Methods

2.1. Participants

We recruited 40 portrait artists; however one artist did not show up for the experiment. The same number of novices were recruited to match artists in age. The sample size was determined by a power analysis with $\alpha = 0.05$, $\beta = 0.2$, an estimated group difference in accuracy 0.05, and an estimated standard deviation 0.08 in each group (i.e., effect size = 0.63) based on a pilot study with 15 artists and 15 novices performing a face matching task (Galmar, Chung, & Hsiao, 2014), assuming a similar effect size across the three tasks if the effect of drawing expertise was task general. In total we had data from 39 artists (17 males, 22 females, mean age = 28.28) and 39 novices (15 males, 24 females, mean age = 29.23) according to their drawing expertise. All participants were Chinese from Hong Kong, whose age ranged from 18 to 66 ($M = 28.76$, $SD = 13.05$). They all had normal or corrected-to-normal vision. The artists were well-trained painters, who had 3 to 40 years ($M = 12.92$) formal drawing training experience, including 1 to 30 years ($M = 8.56$) experience in face-drawing. In contrast, novices did not receive any formal training in drawing (i.e., they were not artists and did not have more than usual exposure to portraits). To further assess the overall drawing level of artists and novices, participants were asked to draw a portrait of a celebrity (either Emma Watson or Barack Obama) given the celebrity's picture in 15 min prior to the experiment. The portraits were then assessed by three professional full-time portrait artists who did not participate in the study and had more than 10-year face drawing experience on a 10-point Likert scale. The average score from the three judges was used to assess participants' drawing expertise. According to this expertise measure, artists outperformed novices in portrait drawing (artists: 5.73/10; novices: 1.67/10), $t(76) = 13.84$, $p < .001$, $d = 3.13$ (equal variances assumed: Levene's test for equality of variances, $F(1, 76) = 0.12$, $p = .73$. $BF_{10} = 1.00 \times 10^{17}$, very strong evidence favoring the alternative hypothesis[1]). The rating scores from the three judges had high internal consistency according to Cronbach's alpha, $\alpha = 0.916$.

In addition, artists and novices performed a verbal and a visuospatial two-back task to assess their verbal and visuospatial working memory capacities (Lau, Ip, Lee, Yeung, & Eskes, 2016). In the two-back task, in each trial, participants were presented with a symbol appearing at a randomly chosen location among 12 possibilities. In the verbal two-back task, participants judged whether the symbol presented in the current trial was the same as the one presented two trials back (while ignoring the symbol location). In the visuospatial two-back task, they judged whether the location of the symbol was the same as the one presented two trials back (regardless of the symbol identity). Artists and novices did not differ significantly in accuracy in either the verbal n-back task (verbal: artists = 73.79%; novices = 77.58%), $t(76) = 0.25$, *n.s.* ($BF_{01} = 1.71$, weak evidence favoring the null hypothesis), or the visuospatial n-back task (artists = 70.84%; novices = 74.83%), $t(76) = 0.176$, *n.s.* ($BF_{01} = 1.68$, weak evidence favoring the null hypothesis), suggesting comparable working memory capacities. The two participant groups also had similar response times (RTs) in the verbal two-back task (artists = 953.86 ms; novices = 884.29 ms), $t(66.51) = 1.87$, *n.s.* ($BF_{10} = 1.47$, weak evidence favoring the alternative hypothesis); however, artists had longer RTs than novices in the visuospatial two-back task (artists = 950.89 ms; novices = 815.63 ms), $t(72.31) = 3.33$, $p = .001$, $d = 0.75$ (Equal variances not assumed: Levene's test for equality of variances, $F(1, 76) = 4.40$, $p = .039$. $BF_{10} = 34.70$, strong evidence favoring the alternative hypothesis).

This study was approval by the Human Research Ethics Committee at the University of Hong Kong (protocol number EA220114). Participants first performed the celebrity portrait drawing, followed by the three experiments described below in sequence (face recognition, face matching, and the part-whole task). They then performed the verbal and visual spatial two-back task in the end.

[1] In the Bayes Factor analysis for *t*-test, scaled information prior with scale $r = 0.707$ was used given that a medium effect size was expected (Rouder, Speckman, Sun, Morey, & Iverson, 2009). For correlations, we used the Jeffreys-Zellner-Siow Cauchy prior with scale $r = 0.354$ (Liang, Paulo, Molina, Clyde, & Berger, 2008). We reported Bayes factors between 1 and 3 as weak, 3 to 10 as positive, 20 to 150 as strong, and > 150 as very strong evidences for the hypothesis (Raftery, 1995).

2.2. Apparatus

The three experiments were conducted using Eprime 2.0 (Psychology Software Tools) combined with a RED-n scientific eye tracker (SensoMotoric Instruments GmbH) on a 17″ laptop with a resolution of 1280 × 768 pixels. Participants' eye movements were recorded with 60 Hz sampling frequency. Smart binocular tracking mode was used, in which both eyes were tracked, and tracking continued when one eye was closed or could not be tracked. Participants' viewing distance was 60 cm. A chinrest was used to reduce head movement. In all experiments, a nine-point calibration and validation procedure was performed before the start of each phase/block (see below for phase/block information). Recalibration was performed if participants' calibration error was more than 0.3° (for the dominant eye) or 0.5° (for the non-dominant eye) horizontal or vertical visual angle. Data from the dominant eye was used for analysis. The background color during the calibration matched the stimulus background in all experiments, although there was a slight mismatch in the face matching experiment due to the experiment requirement of having different lighting conditions across stimuli.[2] Fixations were identified from the raw eye movement data using the software SMI BeGaze (SensoMotoric Instruments GmbH) with its default settings.

2.3. Experiment 1: Face recognition

2.3.1. Materials

The stimuli were 64 frontal-view Chinese face color images; 32 of them were young adult and the other 32 were older adult faces. Half of the faces in each age group were male faces. All faces had a neutral expression and were unfamiliar to the participants. The face stimuli were cropped according to the original face shape with the hair, ears, and neck area removed (see Fig. 1A for examples).

2.3.2. Procedure

The face recognition task consisted of two phases: a study phase and a test phase. In the study phase, each trial started with a central fixation. After detecting participants' central fixation using the eye tracker, a circle appeared around the cross at the center of the screen. The experimenter then pressed a button to present a target face stimulus either in the left visual field (LVF) or the right visual field (RVF) for 3000 ms in a random order (Fig. 2B). In total there were 8 older adult and 8 young adult target faces. Each face stimulus subtended a horizontal and vertical visual angle of 8° x 10° (see Fig. 1A). The center of the face stimulus was 7.9° visual angle away from the center of the screen. Participants were asked to remember all faces presented in the study phase. The test phase started immediately after the study phase.

In the test phase, similarly, in each trial a target face stimulus was presented either in the LVF or the RVF after detecting participants' central fixation. The target face stimuli consisted of 16 old faces (i.e., faces that have been presented in the study phase) and 16 new faces (i.e., faces that have not been shown before); half of the stimuli were male faces (Fig. 1B). Participants judged whether they have seen the face in the study phase by pressing buttons on a response box with both hands. During the experiment, participants performed the face recognition task twice with two different sets of stimuli.

[2] Recent research has suggested that pupil-size change due to luminance change during eye tracking may lead to idiosyncratic systematic errors (e.g., Drewes, Zhu, Hu, & Hu, 2014; Hooge, Hessels, & Nystrum, 2019). In the current study, the eye movement pattern measure using EMHMM (eyes-nose scale; please see the Results section) had an excellent split-half reliability (face recognition, Cronbach's $\alpha = 0.994$; face matching, Cronbach's $\alpha = 0.998$; part-whole task, Cronbach's $\alpha = 0.982$), suggesting that it is a reliable measure despite the potential errors.

2.4. Experiment 2: Face matching

2.4.1. Materials

The materials consisted of 264 pairs of frontal-view color images of 53 young and 57 older Chinese adult face models, with half male and half female face images. The two face images in a pair were matched in sex and age. They could differ in the number of accessories used, ranging from 0 to 3, including use of a wig (i.e., change in hairstyle), use of a hat, and use of glasses. They might also differ in facial expression (two levels: neutral face vs. emotional face, e.g., surprised, happy, etc.) and lighting condition (three levels: normal vs. amber vs. dark; see Fig. 2 for examples). All stimuli were aligned according to the eye height and face midline.

2.4.2. Procedure

Each trial started with a central fixation. After detecting participants' central fixation using the eye tracker, a circle appeared around the cross at the center of the screen. Two face images in a pair (see Materials) were then presented on the left and right side of the screen simultaneously for 3000 ms, followed by a grey pattern mask (Fig. 2). Participants were asked to judge whether the two face images were the same person by pressing corresponding keys on a response box with both hands. This is to avoid any lateralization influence that may be caused by one-hand responses (Mohr, Pulvermüller, & Zaidel, 1994). Each face image subtended a horizontal and vertical visual angle of 8° x 10° (consistent with, e.g., Hsiao & Cottrell, 2009; Chuk et al., 2014; Chuk et al., 2017; Fig. 1). The center of each face was 7.9° visual angle away from the center of the screen. In total, 264 trials were randomly presented in 8 blocks, with 33 trials in each block. Each target face stimulus ($n = 132$) appeared once in a 'same' trial and once in a 'different' trial. Other face stimuli to be compared with the target face stimuli were only presented once in the task. Trials with face stimulus pairs differed in different numbers of accessories (0, 1, 2, or 3) were equally distributed among the 'same' ($n = 132$) and 'different' trials ($n = 132$).

2.5. Experiment 3: Part-whole task

2.5.1. Materials

The stimuli consisted of 24 young Asian male artificial face color images generated using the software FaceGen Modeller (https://facegen.com/modeller.htm), and 36 facial feature images cropped from the 24 face images, including the eyes ($n = 12$), nose (n = 12), and mouth (n = 12; Fig. 3A). All faces had neutral expressions and were unfamiliar to the participants. Each face subtended a horizontal and vertical visual angle of 8° x 10° on the screen under a 60 cm viewing distance. Each pair of eyes subtended a visual angle of 8° x 1.39°; each nose subtended a visual angle of 1.90° x 2.53°, and each mouth subtended a visual angle of 2.78° x 0.76°. Stimuli were presented on a black background.

2.5.2. Procedures

The part-whole task consisted of two phases: a study phase and a test phase (Tanaka & Farah, 1993). In the study phase, each trial started with a central fixation. After detecting participants' central fixation, a circle appeared around the cross at the center of the screen. The experimenter then pressed a button to present the target face stimulus at the center of the screen, together with a corresponding name appearing on top of the screen. The stimulus was presented on the screen for 5000 ms (Fig. 3C). Six face images were presented in the study phase one at a time in a random order, and this presentation was repeated for five times (30 trials in total). The test phase started immediately after the study phase.

The test phase consisted of two conditions: the whole condition and the part condition. Each trial started with a central fixation. After detecting participants' central fixation, the experimenter pressed a button to present two stimuli simultaneously with each at the center of the left-half and the right-half of the screen respectively. In the whole condition, participants were presented with a target face and a foil face

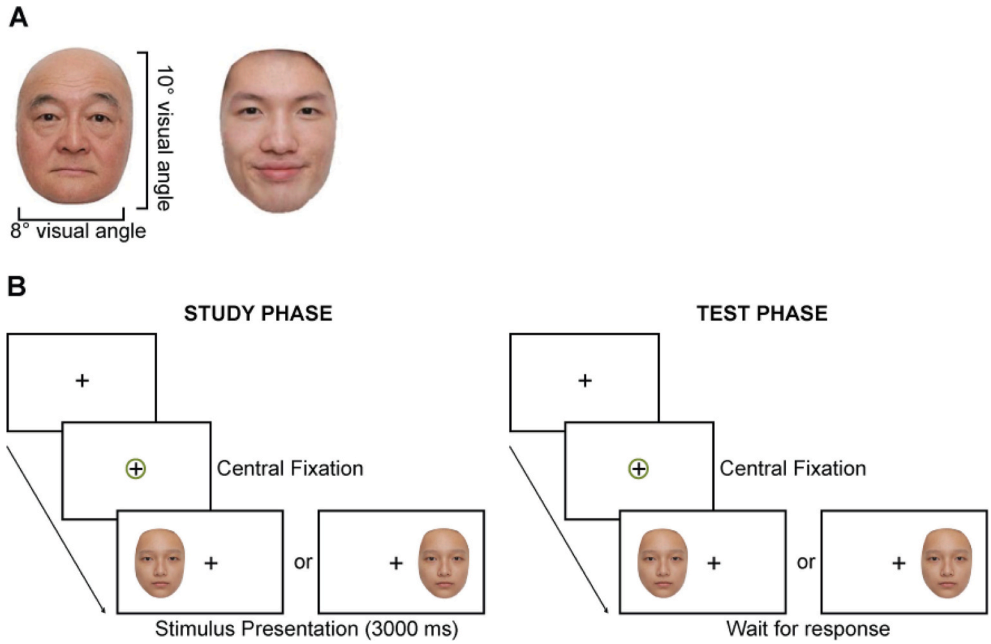


Fig. 1. (A) Examples of face stimuli used in the face recognition task: an older adult face (left), and a young adult face (right). (B) Procedure of the study phase (left) and the test phase (right) of the face recognition task.

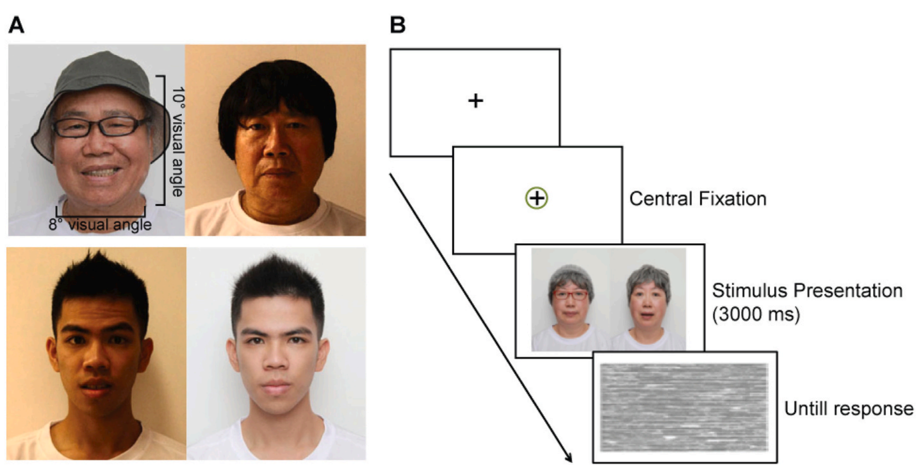


Fig. 2. (A) Examples of face stimuli used in the face matching task: a pair of older adult face stimuli with different accessories (hairstyle, hat, and glasses), facial expressions, and lighting conditions (top), and a pair of young adult face stimuli with no accessory difference but with different lighting conditions and facial expressions (bottom). (B) Procedure of the face matching task.

that differed from the target face in only one facial feature (e.g., the eyes, nose, or mouth), with a question appearing on the top of the screen asking participants to identify a face they saw during the study phase (e.g., Which is Peter?). In the part condition, participants were presented with two images of the same facial feature (e.g., two noses), with a question appearing on the top of the screen asking them to judge which one belonged to a given person/face they saw during the study phase (e.g., Which is Peter's nose? See Fig. 3C). Participants made their judgments by pressing corresponding buttons on a response box with both hands. The whole and part conditions were randomly presented, with 18 trials in each condition (following Joseph & Tanaka, 2002). In the part condition, there were 6 trials for each of the 3 facial features. The facial features that differed/used in the whole and part conditions were matched, and thus the only difference between the two conditions was whether the pair of facial features were presented within a whole face context or in isolation.

2.6. Eye movement data analysis

We used the Eye Movement analysis with Hidden Markov Models (EMHMM) approach (Chuk et al., 2014; http://visal.cs.cityu.edu.hk/research/emhmm/) to analyze the eye gaze location and transition data in the face matching, face recognition, and part-whole tasks. In contrast to other approaches that directly compare eye fixation data between participant groups using predefined regions of interest (ROIs; e.g., Henderson, Williams, & Falk, 2005) or statistical fixation heat maps (e.g., Caldara & Miellet, 2011), this machine learning approach summarizes an individual's eye movement pattern in terms of personalized ROIs (as 2D Gaussian distributions) and transition probabilities among the ROIs using a hidden Markov model (HMM), which is a type of time-series probabilistic model that defines a probability density over eye fixation sequences. Thus, individual differences in both spatial and temporal dimensions of eye movements (i.e., eye gaze locations and transitions) can be better captured using this method. Individual HMMs

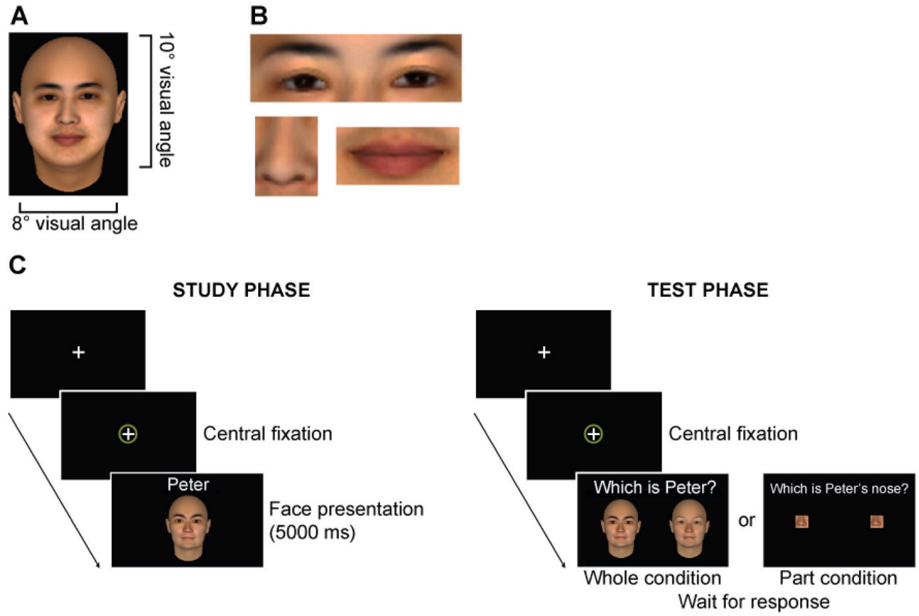


Fig. 3. Examples of (A) a whole face image and (B) facial feature stimuli, including eyes, a nose, and a mouth, used in the part-whole task. (C) Procedure of the study phase (left) and test phase (right) of the part-whole task.

can be clustered using the variational hierarchical expectation maximization (VHEM) algorithm (Coviello, Chan, & Lanckriet, 2014) to discover common eye movement patterns in the population. VHEM clusters the individuals' HMMs based on their probability densities, forming groups with similar distributions of eye fixation sequences, and estimates a representative HMM for each group, which summarizes the common eye gaze pattern for that group. More specifically, the algorithm first initializes each representative HMM with a randomly selected input HMM. Then, in the *E*-step, the similarities (expected log likelihoods) of each input HMM to the representative HMMs are estimated. In the M-step, the input HMMs are clustered according to the similarities. These two steps iterate until convergence. This procedure is typically performed for multiple times, and the result with the highest expected log likelihood is used. The number of strategies (i.e., number of clusters) that the VHEM algorithm discovers can be either pre-determined (Coviello et al., 2014), or automatically estimated from the data using Bayesian methodology (Lan, Liu, Hsiao, Yu, & Chan, submitted). Note that because the HMM defines a probability density over eye fixation sequences, it is possible to calculate the likelihood of an eye fixation sequence being generated from the HMM. Thus, similarities between individual eye movement patterns and the discovered common patterns can be quantitatively assessed by calculating the log-likelihood of the individual eye gaze fixation and transition data being generated by the representative HMMs of the common patterns. This log-likelihood measure provides a quantitative measure of eye movement pattern similarities among individuals, and thus is particularly suitable for the examination of the relationship between eye movement patterns and other cognitive measures in the current study.

In the current analysis, similar to our previous studies, we included only the eye fixations on the face area and removed the fixations on the hair, ears, neck, and background using a predefined area of interest (face recognition, two ellipses centered at (320, 377.3) and (960, 337.5) respectively with 245 and 207.5 pixels as the horizontal and vertical radius; face matching, two rectangles defined by $160 \leq x \leq 510$, $810 \leq x \leq 1140$, and $140 \leq y \leq 500$; part-whole task, two rectangles defined by $140 \leq x \leq 515$, $770 \leq x \leq 1145$, and $85 \leq y \leq 630$. E.g., Chuk et al., 2017; Zhang et al., 2019). For each task, after we trained one HMM to summarize a participant's eye movement pattern, we clustered the individual HMMs to discover two common eye movement patterns from the participants using the VHEM algorithm, following previous works (Chan et al., 2018; Chuk et al., 2014; Chuk, Chan, & Hsiao, 2017) that set the number of clusters to 2. Note that results from a new VBHEM methodology that uses Bayesian methods to determine the optimal number of clusters (Lan et al., submitted) also suggested two clusters for the data in all three tasks reported here. We then used the representative HMMs of the two common patterns to obtain quantitative measures of eye movement pattern similarity. When training individual HMMs, EMHMM uses the Bayesian method to determine automatically the optimal number of ROIs for each individual from a pre-set range (here we used 1 to 6 ROIs). Each model with a specific number of ROI was trained for 100 times, and the model with the highest data log-likelihood within the pre-set range was used in the analysis. When clustering individual HMMs to generate representative HMMs of common patterns, following our previous studies (e.g., Chuk et al., 2017; Chan et al., 2018; Zhang et al., 2019), we used the median number of ROIs among the individual models. Accordingly, we used 6 ROIs for generating representative models for the face matching task, 5 ROIs for face recognition task, and 5 ROIs for the part-whole task. The clustering algorithm was repeated 100 times with different initializations, and the result with the highest data log-likelihood was used. The data that support the findings of this study are openly available at http://doi.org/10.17605/OSF.IO/ENJMC.

3. Results

3.1. Experiment 1: Face recognition

The design consisted of one between-subject variable expertise (artist vs. novice). The dependent variable was recognition performance as measured in d' and correct response time (RT).

3.1.1. Performance on the task

As compared with novices, portrait artists ($M = 1.67$, $SD = 0.56$) had higher recognition sensitivity as measured in d' ($M = 1.36$, $SD = 0.63$), $t(74) = 2.27$, $p = .026$, $d = 0.52$ (Equal variances assumed: Levene's test for equality of variances, $F(1, 74) = 0.47$, $p = .496$. $BF_{10} = 2.99$, weak

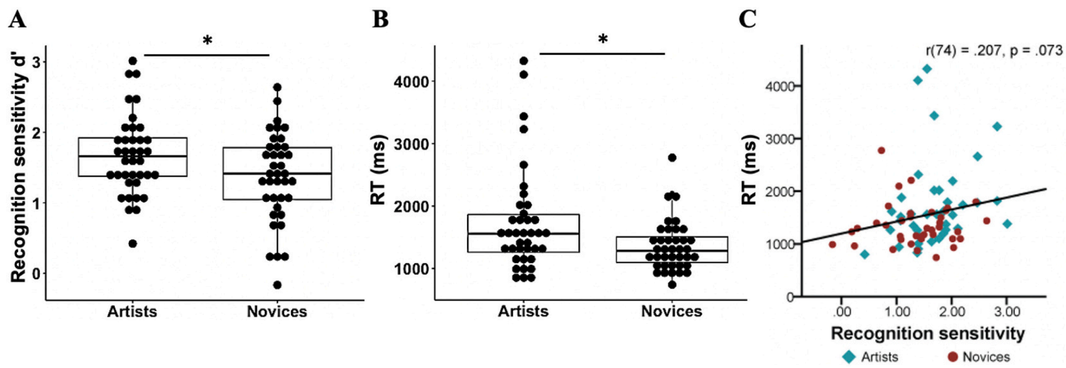


Fig. 4. (A, B) Face artists had higher recognition sensitivity d' but longer RT than novices in face recognition (*$p < .05$). (C) Marginal correlation between face recognition sensitivity d' and RT.

evidence favoring the alternative hypothesis; Fig. 4A), but a longer RT, $t(53.20) = 2.56$, $p = .013$, d = 0.59 (Equal variances not assumed: Levene's test for equality of variances, $F(1, 74) = 7.59$, $p = .007$. $BF_{10} = 5.44$, positive evidence favoring the alternative hypothesis; Fig. 4B).[3] These findings indicate that artists had better recognition sensitivity but also took longer to respond in the face recognition task, suggesting a speed-accuracy tradeoff. Consistent with this speculation, we observed a marginal positive correlation between face recognition sensitivity and RT among participants, $r(74) = 0.21$, $p = .073$, 95% CI [−0.02, 0.43] (Fig. 4C; $BF_{10} = 1.14$, weak evidence favoring the alternative hypothesis), and this correlation was significant when we used a more robust method, the percentage bend correlation (Pernet, Wilcox, & Rousselet, 2013), $r(74) = 0.24$, $p = .041$, 95% CI [0.01, 0.45]. This result suggested that artists had better recognition sensitivity than novices at the expense of their RT.[4]

3.1.2. Gaze patterns

In eye movement data analysis, individual HMMs were generated using test phase eye gaze location and transition data. These HMMs were clustered according to similarities to discover two common patterns for face recognition. The median number of ROIs across individual HMMs was 5, and thus 5 ROIs were used to generate the representative models of the common patterns. Fig. 5 shows the resulting representative HMMs of the two common patterns. For the current purposes, we refer to them as the eyes-focused and nose-focused patterns for face recognition. The eyes-focused pattern (Fig. 5A) consisted of ROIs that centered around the two eyes. Most initial fixations occurred at the eye region (red ROI, 96%). From the first fixation, participants most likely stayed within the eye region, either moving to the right eye (pink ROI, 28%), the left eye (cyan ROI, 22%), or to the eye region in general (blue ROI, 18%); occasionally they moved to areas across the face center (green ROI, 32%). The fixation heat map shows that most fixations concentrated on the individual eyes. The second representative model, labeled the nose-focused pattern, consisted of ROIs that centered on the nose bridge (Fig. 5B). The fixation heat map shows that most fixations concentrated on the bridge of the nose. The representative HMMs of the two common patterns significantly differed from each other: data from participants adopting the eyes-focused pattern were significantly more likely to be generated by the eyes-focused HMM than the nose-focused one, $t(38) = 10.15$, $p < .001$, $d = 1.63$, and data from those adopting the nose-focused pattern were more likely to be generated by the nose-focused HMM than the eyes-focused one, $t(35) = 8.76$, $p < .001$, $d = 1.46$. In summary, in the eyes-focused pattern for face recognition, participants' eye fixations sometimes switched between the face center and the midpoint between the two eyes, and sometimes switched between the left and right eyes. In contrast, in the nose-focused pattern, participants' fixations tended to stay around the center and midline of the face with less clear transitions among facial features.

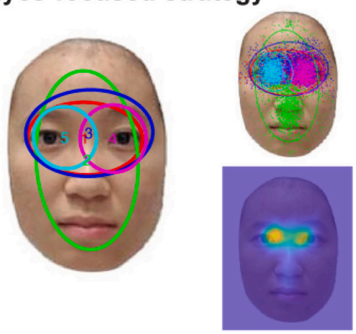
To quantify a participant's eye movement pattern along the continuum between the eyes-focused and nose-focused pattern, we defined the eyes-nose scale as $(L_{eye} - L_{nose})/(|L_{eye}| + |L_{nose}|)$, where L_{eye} is the log likelihood of the eye movement pattern being generated by the representative HMM of the eyes-focused pattern, and L_{nose} is the log likelihood of the eye movement pattern being generated by the representative HMM of the nose-focused pattern (Chan et al., 2018). The larger the eyes-nose scale, the more similar the pattern is to the eyes-focused pattern, and vice versa for the nose-focused pattern. To see whether artists and novices differed in their eye movement patterns, we compared the eyes-nose scale of the artists and novices.[5] Artists and Novices did not differ significantly in eyes-nose scale for face recognition, $t(73) = 0.04$, $p = .97$ (Equal variances assumed: Levene's test for equality of variances, $F(1, 73) = 1.95$, $p = .166$. $BF_{01} = 3.22$, positive evidence favoring the null hypothesis), suggesting that they did not differ in eye movement pattern in face recognition. Note that artists and novices also did not differ in average fixation duration, $t(73) = -1.31$, $p = .20$ (Equal variances assumed: Levene's test for equality of variances, $F(1, 73) = 0.46$, $p = .502$. $BF_{01} = 1.48$, weak evidence favoring the null hypothesis); however, artists made significantly more fixations per trials than novices, $t(73) = 2.39$, $p = .019$, d = 0.55 (Equal variances assumed: Levene's test for equality of variances, $F(1, 73) = 1.95$, $p = .167$; $BF_{10} = 3.83$, positive evidence favoring the alternative hypothesis). This result was consistent with the finding that artists had significant longer RTs than novices.

[3] In the face recognition task, two participants had missing behavioral data and three participants had missing eye movement data due to technical problem.

[4] Note that we identified two outliers in the RT data when using three times of the standard deviation above and below the mean as the criterion. After removing the outliers, the correlation between d' and RT was significant, $r(72) = 0.28$, $p = .014$. When the correlation was conducted separately for artists and novices, d' was positively correlated with RT among artists, $r(34) = 0.43$, $p = .009$, 95% CI [0.12, 0.79], but not among novices, $r(36) = 0.00$, $p = .10$, 95% CI [−0.59, 0.59]. We then examined moderation effect of group (artist vs. novice) on the relation between d' and RT. In the first step, recognition sensitivity (d') and group were entered in the regression analysis. Next, the interaction term between recognition sensitivity and group was entered, and it accounted for a significant increase in variance in RT, $\Delta R^2 = 0.07$, $\Delta F(1, 70) = 5.57$, b = −444.96, $p = .021$. The overall model explained approximately 17.49% of the variance in RT, $F(3, 70) = 4.95$, $p = .004$. This result suggested a speed-accuracy tradeoff in artists but not in novices.

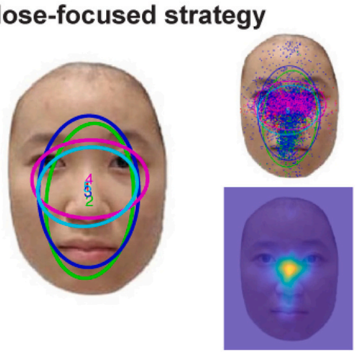
[5] In a separate analysis, we performed the clustering only on novices' data and used the resulting representative models to quantify participants' eye movement patterns. Similar results were obtained in all three face tasks.

A. Eyes-focused strategy



	To red	To green	To blue	To pink	To cyan
Prior	.96	.01	.03	.00	.00
From red	.00	.32	.18	.28	.22
From green	.01	.39	.59	.01	.00
From blue	.00	.39	.60	.00	.00
From pink	.01	.52	.09	.00	.38
From cyan	.00	.45	.00	.55	.00

B. Nose-focused strategy



	To red	To green	To blue	To pink	To cyan
Prior	.90	.03	.00	.00	.08
From red	.00	.50	.06	.44	.00
From green	.00	.09	.69	.21	.00
From blue	.00	.23	.67	.11	.00
From pink	.00	.35	.48	.18	.00
From cyan	.00	.51	.04	.45	.00

Fig. 5. The two representative eye movement patterns in face recognition task derived by clustering using EMHMM. The top image shows the representative eyes-focused pattern, and the bottom image shows the representative nose-focused pattern.

3.1.3. *Relation between gaze patterns and performance*

We also examined whether the two representative eye movement patterns discovered in face recognition were associated with different face recognition performance. We found that participants who adopted the eyes-focused pattern had higher d' than those using the nose-focused pattern, $t(72) = -2.30$, $p = .025$, d $= -0.53$ (Equal variances assumed: Levene's test for equality of variances, $F(1,72) = 0.93$, $p = .338$. $BF_{10} = 3.19$, positive evidence favoring the alternative hypothesis) but did not differ significantly from those using the nose-focused pattern in RT, $t(72) = -0.94$, $p = .35$ (Equal variances assumed: Levene's test for equality of variances, $F(1,72) = 0.79$, $p = .378$. $BF_{01} = 3.05$, positive evidence favoring the null hypothesis;). This finding was consistent with previous studies (e.g., Chuk et al., 2017; Chuk, Crookes, Hayward, Chan, & Hsiao, 2017; Chan et al., 2018). Interestingly, we observed a significant linear relationship between eyes-nose scale and recognition sensitivity d', $R^2 = 0.11$, $F(1, 72) = 9.06$, $p = .004$, $\beta = -4.37$ (Fig. 6A; $BF_{10} = 10.43$, positive evidence favoring the alternative hypothesis). Since previous studies have suggested that eye movement patterns that involve looking at both the nose and the eyes are optimal for face recognition (e.g., Chuk, Crookes, et al., 2017), we performed an exploratory analysis to examine the possibility of a curvilinear relationship between eyes-nose scale and recognition performance. The results showed a significant quadratic relationship between eyes-nose scale and recognition sensitivity d', $R^2 = 0.20$, $F(2, 71) = 9.08$, $p < .001$, $\beta 1 = -3.52$, $\beta 2 = -80.12$ (Fig. 6B), and it accounted for more variance than the linear model, $\Delta R^2 = 0.09$, $p = .006$.[6] This result indicated that those who used either a highly eyes-focused or a highly nose-focused pattern tended to have lower recognition sensitivity, while those whose eye movement patterns were a mixture of the two patterns tended to have higher recognition sensitivity.

[6] Note that adding a cubic term at the next step did not significantly account for more variance, $\Delta R^2 = 0.008$, $p = .39$.

3.2. *Experiment 2: Face matching*

The design consisted of one between-subject variable expertise (artist vs. novice). The dependent variable was recognition performance as measured in d' and correct response time (RT).

3.2.1. *Performance on the task*

When we compared artists' and novices' discrimination sensitivity as measured in d', portrait artists had better d' than novices, $t(76) = 4.05$, $p < .001$, d $= 0.92$ (Equal variances assumed: Levene's test for equality of variances, $F(1, 76) = 0.04$, $p = .852$. $BF_{10} = 261.52$, very strong evidence favoring the alternative hypothesis; Fig. 7A). In correct RT, no significant difference was observed between the two groups, $t(76) = 0.16$, $p = .88$ (Equal variances assumed: Levene's test for equality of variances, $F(1,76) = 1.08$, $p = .302$. $BF_{01} = 3.24$, positive evidence favoring the null hypothesis).

3.2.2. *Relation between expertise and performance*

Participants' face drawing rating was positively correlated with face matching discrimination sensitivity d', $r(76) = 0.52$, $p < .001$, 95% CI [0.32, 0.71] (Fig. 7B; $BF_{10} = 1.65 \times 10^4$, very strong evidence favoring the alternative hypothesis): the better the face drawing rating, the better the face matching performance. A similar result was obtained using the percentage bend correlation $r(76) = 0.52$, $p < .001$, 95% CI [0.34, 0.67]. When we examined the correlation between drawing rating and face matching performance in the two groups separately, it was significant among artists, $r(39) = 0.385$, $p = .016$, 95% CI [0.05, 0.45] ($BF_{10} = 4.60$, positive evidence favoring the alternative hypothesis), and was marginal among novices, $r(39) = 0.277$, $p = .087$, 95% CI [−0.022, 0.31] ($BF_{10} = 1.26$, weak evidence favoring the alternative hypothesis).

3.2.3. *Gaze patterns*

Through clustering individual HMMs of eye faze location and transition data according to similarities, we discovered two common eye movement patterns, as shown in Fig. 8. For the current purposes, we

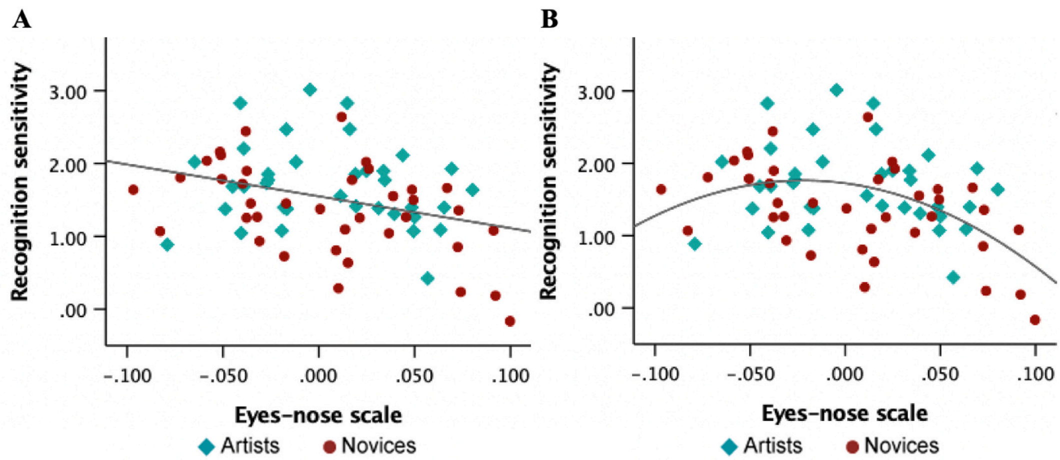


Fig. 6. (A) A linear relationship, and (B) a quadratic relationship between eyes-nose scale and recognition sensitivity (d') in the face recognition task.

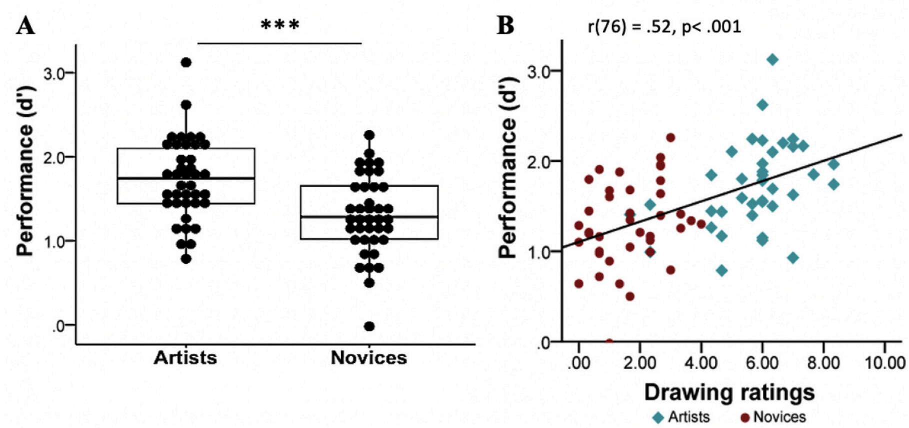
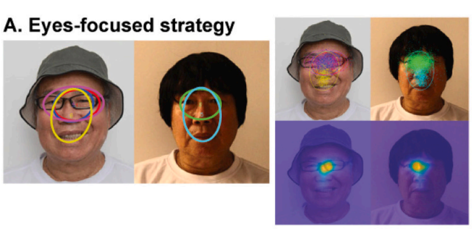
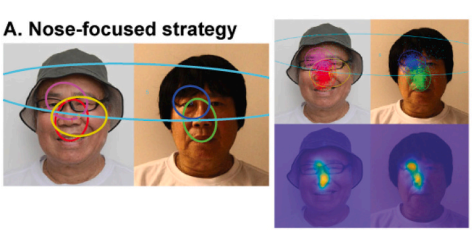


Fig. 7. (A) Mean performance (d') of artists and novices in the face matching task (***p < .001). (B) Correlations between participants' drawing ratings assessed by three professional portrait artists and their performance (d') in the face matching task.



	To red	To green	To blue	To pink	To cyan	To Yellow
Prior	.71	.24	.00	.02	.02	.02
From red	.00	.74	.01	.16	.05	.04
From green	.02	.19	.65	.00	.10	.04
From blue	.00	.49	.01	.21	.16	.14
From pink	.01	.68	.00	.14	.09	.08
From cyan	.00	.10	.01	.00	.26	.63
From yellow	.00	.00	.02	.11	.65	.21



	To red	To green	To blue	To pink	To cyan	To Yellow
Prior	.31	.09	.29	.14	.01	.16
From red	.19	.74	.00	.02	.00	.04
From green	.72	.25	.02	.00	.00	.01
From blue	.02	.04	.15	.78	.00	.00
From pink	.04	.04	.73	.19	.00	.00
From cyan	.00	.00	.01	.04	.95	.00
From yellow	.39	.59	.00	.00	.00	.02

Fig. 8. The two representative eye movement patterns discovered in the face matching task through clustering using EMHMM: Eyes-focused and nose-focused patterns for face matching. The ellipses on the images show regions of interest (ROIs) as 2-D Gaussian distributions (solid lines mark 2 standard deviations away from the mean). The small image on the top right shows the distribution of actual fixations (subsampled for better visualization) that belong to each ROI, whereas the small image on the bottom right shows a corresponding fixation heat map. The table shows the transition probabilities among the ROIs. The prior in the table indicates the probability that a fixation sequence starts from each ROI. Note that the big cyan ROI in the nose-focused pattern captures outlier fixations that do not fit well with the other ROIs on the face. Please see the Eye movement data analysis section in the Methods for details about how number of ROIs was determined. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

refer to them as the eyes-focused and nose-focused patterns for face matching. In the representative HMM of the eyes-focused pattern, the first fixation was most likely to be on the eye region of the left face (red ROI, 71%), followed by the eye region of the right face (green ROI, 24%). After the first fixation, they were most likely to switch between the eye region of the two faces (red to green ROI, 74%; green to blue ROI, 65%; and blue and pink ROIs), with a smaller probability of looking at the center/nose region of the two faces (yellow and cyan ROIs). In contrast, in the representative HMM of the nose-focused pattern, the first fixation was most likely to be on the nose of the left face (red ROI, 31%), and afterwards to switch between the nose region of the two faces (red to green, 74%; green to red, 72%). Occasionally, the first fixation started from the left eye of the right face (blue ROI, 29%), which was most likely followed by another fixation to the eye region of the left face (blue to pink, 73%). Note that the cyan ROI is for some outlier first fixations, since there is 1% probability of starting there, and zero probability to transition to the cyan ROI from other ROIs. The representative HMMs of the two patterns significantly differed from each other according to Kullback-Leibler (KL) divergence approximation using data log-likelihoods (Chuk et al., 2014): data from participants adopting the eyes-focused pattern were significantly more likely to be generated by the eyes-focused representative HMM than the nose-focused one, $t(47) = 12.11$, $p < .001$, $d = 1.75$, and data from those adopting the nose-focused pattern were more likely to be generated by the nose-focused representative HMM than the eyes-focused one, $t(29) = 8.50$, $p < .001$, $d = 1.55$. In summary, in the eyes-focused pattern for face matching, participants mainly switched between the eye regions of the two faces, and occasionally looked at and then switched between the centers of the two faces. In contrast, in the nose-focused pattern for face matching, participants mainly switched between the centers of the two faces, with a small probability to look at and then switched between the eye regions of the two faces.

Following the previous analysis conducted for the face recognition task, eyes-nose scale was calculated to quantify the similarity of a participant's eye movement pattern along the dimension between the two representative eye movement patterns. A *t*-test showed that artists' eye movement patterns were more eyes-focused than novices, $t(67.88) = 2.40$, $p = .018$ (Equal variances not assumed: Levene's test for equality of variances, $F(1, 67.88) = 8.978$, $p = .004$. $BF_{10} = 4.05$, positive evidence favoring the alternative hypothesis. Fig. 9A).

3.2.4. Relation between expertise and gaze patterns

Participants' face drawing rating was positively correlated with eyes-nose scale, $r(76) = 0.31$, $p = .006$, 95% CI [0.09, 0.52] (Fig. 9B; $BF_{10} = 8.16$, positive evidence favoring the alternative hypothesis; percentage bend correlation: $r(76) = 0.28$, $p = .014$, 95% CI [0.05, 0.46]): the higher the face drawing rating, the more eyes-focused the eye movement pattern. Note however that when we examined this correlation in the two groups separately, it was not significant in either the artist group, $r(39) = 0.088$, $p = .596$, 95% CI [−0.19, 0.32] ($BF_{01} = 2.55$, weak evidence favoring the null hypothesis), or the novice group, $r(39) = 0.288$, $p = .163$, 95% CI [−0.04, 0.23] ($BF_{01} = 1.22$, weak evidence favoring the null hypothesis), suggesting that the correlation was mainly driven by group difference. Artists and novices did not differ in average number of fixations per trial, $t(76) = 1.30$, $p = .20$ (Equal variances assumed: Levene's test for equality of variances, $F(1,76) = 0.27$, $p = .603$. $BF_{01} = 1.53$, weak evidence favoring the null hypothesis), or average fixation duration, $t(76) = -1.53$, $p = .13$ (Equal variances assumed: Levene's test for equality of variances, $F(1,76) = 0.343$, $p = .560$. $BF_{01} = 1.14$, weak evidence favoring the null hypothesis).

3.2.5. Relation between gaze patterns and performance

To assess whether the two representative eye movement patterns resulted in different face matching performance, two *t*-tests were conducted: Participants who adopted the eyes-focused and nose-focused patterns did not differ in performance, $t(76) = 0.56$, $p = .58$ (Equal variances assumed: Levene's test for equality of variances, $F(1,76) = 0.04$, $p = .839$. $BF_{01} = 2.77$, weak evidence favoring the null hypothesis), or RT, $t(76) = 1.10$, $p = .275$ (Equal variances assumed: Levene's test for equality of variances, $F(1,76) = 3.56$, $p = .063$. $BF_{01} = 1.85$, weak evidence favoring the null hypothesis), in face matching. There was no significant linear relationship between eyes-nose scale and face matching task performance, $r(76) = 0.16$, $p = .17$ ($BF_{01} = 2.23$, weak evidence favoring the null hypothesis). To match the analysis we conducted in Experiment 1, we performed an exploratory analysis to examine the possibility of a curvilinear relationship between eyes-nose scale and face matching performance. A significant quadratic relationship was observed, $R^2 = 0.15$, $F(2, 75) = 6.67$, $p = .002$, $\beta1 = -0.08$, $\beta2 = -150.58$ (Fig. 10A). Nevertheless, this quadratic relationship was not significant after we removed the outlier data point at the bottom left corner (Fig. 10B), $R^2 = 0.05$, $F(2, 74) = 1.95$, $p = .15$, $\beta1 = -0.31$, $\beta2 = -118.30$. Thus, although artists adopted a more eyes-focused eye movement pattern than novices, their higher accuracy in face matching as compared with novices could not be accounted for by this difference in eye movement pattern. Instead, it could be predicted by drawing rating, suggesting that this advantage may be related to their better ability in extracting identity-invariant information between two faces.

3.2.6. Comparison between face recognition and face matching

The above results showed that eye movement pattern predicted performance in face recognition but not that in face matching. Here we conducted a moderation analysis to examine whether the task nature (face matching vs. face recognition) significantly moderated the relationship between eye movement pattern (eyes-nose scale) and task performance d'. In the first step of the regression analysis, task nature and eyes-nose scale were entered. In the second step of the regression analysis, the interaction term between task nature and eyes-nose scale was entered, and it explained a significant increase in variance in performance, $\Delta R^2 = 0.047$, $F(1, 148) = 7.56$, $p = .007$ ($BF_{10} = 4.79$, positive evidence favoring the alternative hypothesis). Thus, task nature was a significant moderator of the relationship between eye movement pattern and task performance. In the face-matching task, eyes-nose scale did not significantly predict performance, $F(1,76) = 1.50$, $p = .225$, $R^2 = 0.006$ ($BF_{01} = 3.78$, positive evidence favoring the null hypothesis), whereas in the face recognition task, eyes-nose scale was a significant predictor of performance, $F(1,72) = 9.06$, $p = .004$, $R^2 = 0.099$ ($BF_{10} = 6.29$, positive evidence favoring the alternative hypothesis).

3.3. Experiment 3: Part-whole task

The design consisted of one between-subject variable expertise (artist vs. novice) and one within-subject variable condition (whole vs. part). The dependent variables were the accuracy and correct RT of the judgments. ANOVA was used.

3.3.1. Performance on the task

In the accuracy data, a significant main effect of condition was found, $F(1, 76) = 170.41$, $p < .001$, $\eta^2 = 0.69$ ($BF_{10} = 9.15 \times 10^{18}$, very strong evidence favoring the alternative hypothesis): participants performed better in the whole condition than the part condition. This effect did not interact with expertise, and there was no main effect of expertise, suggesting that portrait artists and novices did not differ in the part-whole effect. In RT, a significant main effect of condition, $F(1, 76) = 11.98$, $p = .001$, $\eta^2 = 0.14$ ($BF_{10} = 20.02$, strong evidence favoring the alternative hypothesis), and a significant main effect of expertise, $F(1, 76) = 14.90$, $p < .001$, $\eta^2 = 0.16$ ($BF_{10} = 122.15$, strong evidence favoring the alternative hypothesis), were observed: participants responded faster in the part condition than the whole condition, and artists had longer RTs than novices (Table 1). In addition, there was a significant interaction between expertise and condition, $F(1, 76) = 4.40$, $p = .039$, $\eta^2 = 0.06$ ($BF_{Inclusion} = 4.37$, positive evidence favoring the alternative hypothesis): artists had a larger condition effect than novices. Note however this interaction effect may be due to a significantly longer RT in artists than in novices in general (i.e., the main effect of expertise in RT). To rule out this possibility, we normalized individual differences in overall accuracy and RT by assessing the normalized whole-part difference as the performance difference between the whole and part conditions divided by their sum (in accuracy: $\frac{whole-part}{whole+part}$; in RT: $-\left(\frac{whole-part}{whole+part}\right)$). A positive whole-part difference indicated better performance in the whole condition

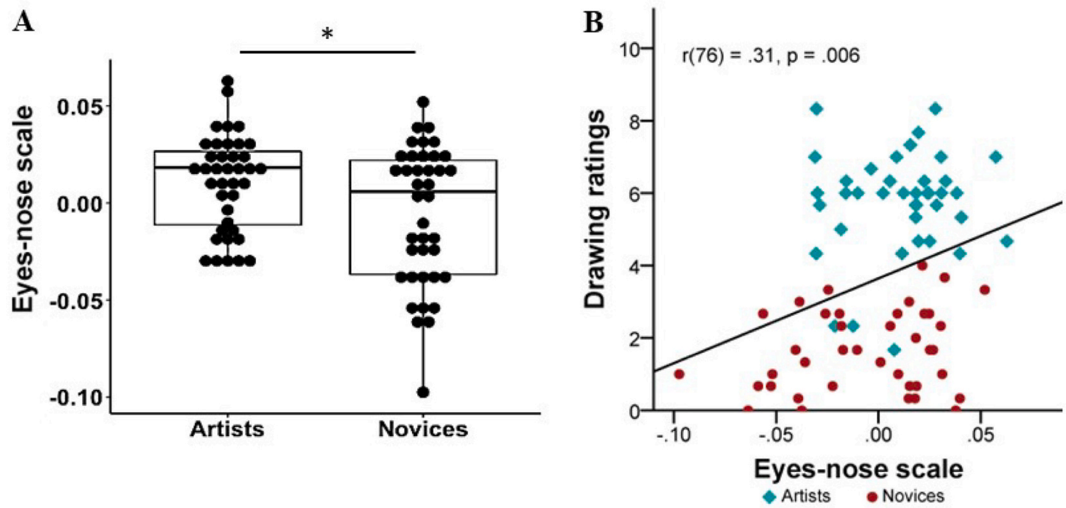


Fig. 9. (A) Eyes-nose scale of artists and novices in the face matching task (*$p < .05$). (B) the correlation between participants' drawing ratings assessed by three professional portrait artists and eyes-nose scale in the face matching task.

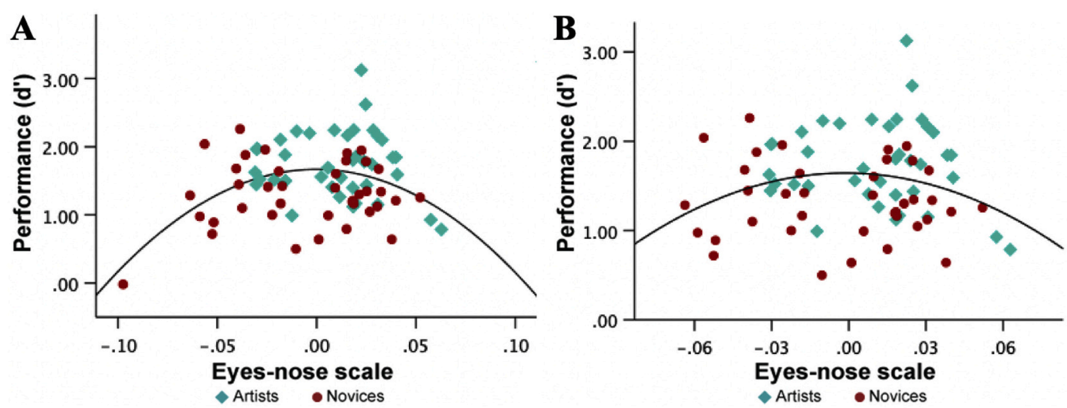


Fig. 10. (A) Quadratic regression between eyes-nose scale and face matching task performance. (B) This quadratic relationship was not significant after removing the outlier data point at the bottom left corner.

Table 1
Accuracy and correct RT of the portrait artists and novices in the part-whole task.

	Artists		Novices	
	M	SD	M	SD
Part condition accuracy	0.68	0.13	0.64	0.10
Whole condition accuracy	0.84	0.11	0.81	0.09
Part condition correct RT (ms)	4102.09	1194.72	3239.57	1162.40
Whole condition correct RT (ms)	4820.04	1868.12	3415.74	1326.24

relative to the part condition, whereas a negative whole-part difference indicated better performance in the part condition relative to the whole condition. The results showed that artists and novices did not differ in the normalized whole-part difference as shown in either the accuracy, $t(76) = -0.46$, $p = .65$ (Equal variances assumed: Levene's test for equality of variances, $F(1,76) = 2.35$, $p = .129$. $BF_{01} = 2.98$, weak evidence favoring the null hypothesis. Fig. 11A) or RT, $t(76) = -1.40$, $p = .17$ (Equal variances assumed: Levene's test for equality of variances, $F(1,76) = 0.69$, $p = .408$. $BF_{01} = 1.35$, weak evidence favoring the null hypothesis; Fig. 11B).

3.3.2. Gaze patterns

We have also examined participants' eye movement patterns during the whole condition of the test phase using the EMHMM method with a between-subject variable expertise (artist vs. novice). For the eye movement analysis, we learned individual HMMs from the whole condition of the test phase in the part-whole task. Similar to the face matching and face recognition tasks, we clustered these individual HMMs into two groups to discover two common patterns. As shown in Fig. 12, for the current purposes, we refer to them as the eyes-focused and nose-focused eye movement patterns for the part-whole task (Fig. 12). In the eye-focused pattern, the first fixation was most likely to be on the eye region of the left face (red ROI, 57%) or the eyes region of the right face (green ROI, 38%). Afterwards, they tended to switch between the eye regions of the two faces: The next fixation from the eye region of the left face had a high probability to move to the eye region of the right face (green ROI, 55%) or stay at the same region (24%). Similarly, the next fixation from the eye region of the right face had a high probability to move to the eye region of the left face (red ROI, 53%) or stay at the same region (27%). In contrast, in the nose-focused pattern, the first fixation was most likely to be on the nose region of the left face (red ROI, 53%). The next fixation from this region had a high probability to move to the nose region of the right face (green ROI, 60%) or stay at the same region (19%). The representative HMMs of the two patterns significantly differed from each other: data from participants adopting the eyes-focused pattern were significantly more likely to be generated by the eyes-focused HMM than the nose-focused one, $t(26)$

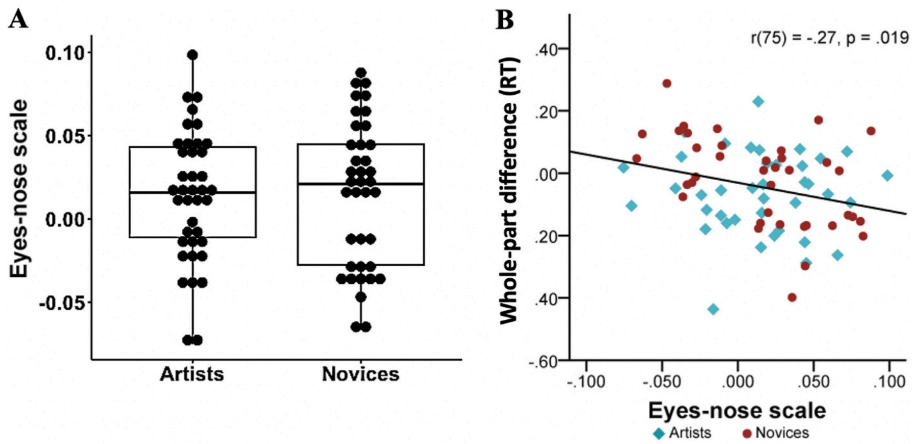
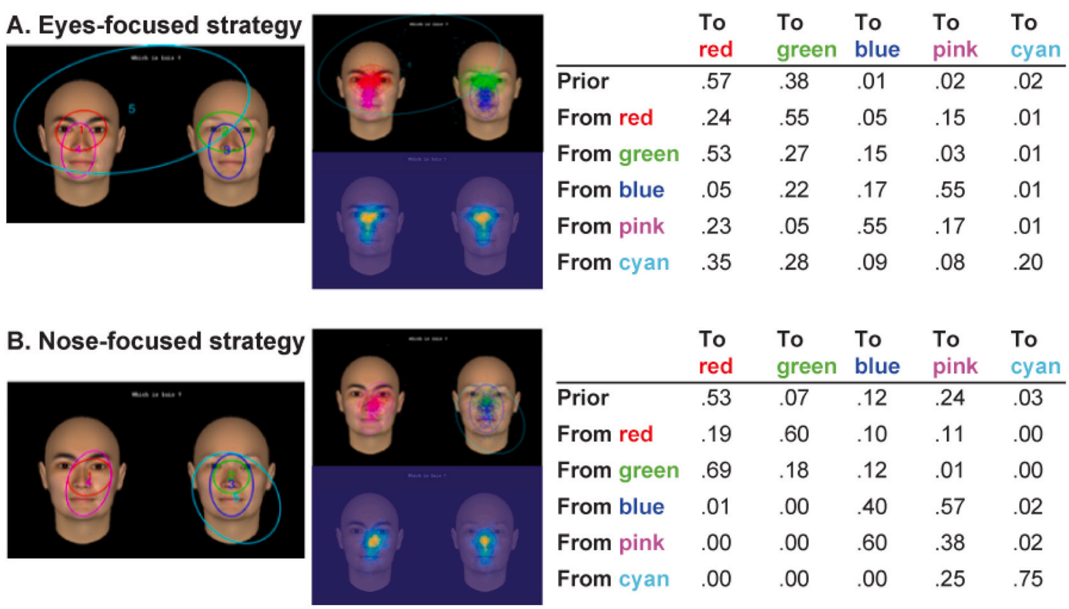


Fig. 11. Part-whole effect (whole-part difference) using the normalized measure in accuracy (A) and in RT (B). Artists and novices did not differ in the part-whole effect.



	To red	To green	To blue	To pink	To cyan
Prior	.57	.38	.01	.02	.02
From red	.24	.55	.05	.15	.01
From green	.53	.27	.15	.03	.01
From blue	.05	.22	.17	.55	.01
From pink	.23	.05	.55	.17	.01
From cyan	.35	.28	.09	.08	.20

	To red	To green	To blue	To pink	To cyan
Prior	.53	.07	.12	.24	.03
From red	.19	.60	.10	.11	.00
From green	.69	.18	.12	.01	.00
From blue	.01	.00	.40	.57	.02
From pink	.00	.00	.60	.38	.02
From cyan	.00	.00	.00	.25	.75

Fig. 12. Eyes-focused and nose-focused eye movement patterns observed in the whole condition of the test phase in the part-whole task.

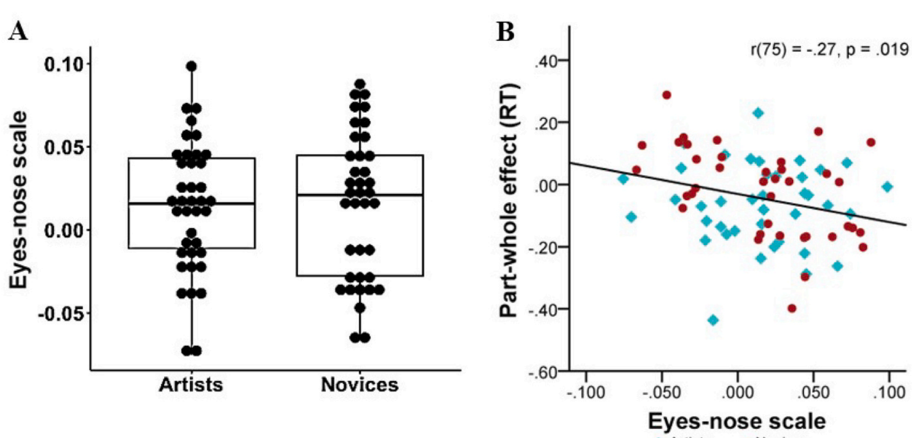


Fig. 13. (A) Portrait artists and novices did not differ in eye movement pattern (as measured in eyes-nose scale using EMHMM) for viewing whole-face stimuli in the test phase of the part-whole task. (B) Participants' eye movement pattern similarity to the eyes-focused pattern (as measured in eyes-nose scale) was correlated with the whole-part difference in RT in the part-whole task: the more eyes-focused the pattern, the longer the RT in the whole condition relative to the part condition.

$= 7.76$, $p < .001$, $d = 1.49$, and data from those adopting the nose-focused pattern were more likely to be generated by the nose-focused HMM than the eyes-focused one, $t(49) = 10.90$, $p < .001$, $d = 1.54$. In summary, in the eyes-focused pattern for the whole-condition in the part-whole task, participants mainly switched between the eye regions of the two faces, and occasionally looked at and switched between the centers of the two faces. In contrast, the nose-focused pattern involved mainly gaze transitions between the centers of the two faces.

Similarly, we calculated the eyes-nose scale of each participant's eye gaze location and transition data in the part-whole task. Artists and novices did not differ significantly in eyes-nose scale, $t(75) = -0.13$, $p = .90$ (Equal variances assumed: Levene's test for equality of variances, $F(1,75) = 1.33$, $p = .253$. $BF_{01} = 3.23$, positive evidence favoring the null hypothesis. Fig. 13A).[7] Note that artists and novices also did not differ in average fixation duration, $t(75) = -0.87$, $p = 39$ (Equal variances assumed: Levene's test for equality of variances, $F(1,75) = 0.07$, $p = .792$. $BF_{01} = 2.30$, weak evidence favoring the null hypothesis). However, artists made significantly more fixations in a trial than novices, $t(75) = 2.88$, $p = .005$, $d = 0.67$ (Equal variances assumed: Levene's test for equality of variances, $F(1,75) = 2.73$, $p = .102$. $BF_{10} = 11.20$, positive evidence favoring the alternative hypothesis), consistent with the finding that they had longer RTs than novices.

3.3.3. Relation between expertise and gaze patterns

We then examined whether the participants using the two eye movement patterns when viewing whole-face stimuli during the test phase differed in the (normalized) whole-part difference. We found that participants who adopted the eyes-focused pattern had a significantly more negative whole-part difference in RT than those using the nose-focused pattern, $t(75) = -2.23$, $p = .029$, $d = -0.53$ (Equal variances assumed: Levene's test for equality of variances, $F(1, 75) = 0.01$, $p = .940$. $BF_{10} = 2.82$, weak evidence favoring the alternative hypothesis;); this effect was not observed in accuracy, $t(75) = -0.71$, $p = .48$ (Equal variances assumed: Levene's test for equality of variances, $F(1,75) = 1.02$, $p = .316$. $BF_{01} = 2.49$, weak evidence favoring the null hypothesis). In addition, participants' eyes-nose scale was correlated with their whole-part difference in the RT, $r(75) = -0.27$, $p = .019$, 95% CI [−0.49, −0.05] ($BF_{10} = 3.35$, positive evidence favoring the alternative hypothesis; Fig. 13B. Percentage bend correlation, $r(75) = -0.27$, $p = .015$, 95% CI [−0.47, −0.05]): the more eyes-focused the pattern, the more negative the whole-part difference in correct RT. No similar correlation was found in accuracy. This result suggested that the eyes-focused eye movement pattern was associated with longer RT in the whole condition relative to the part condition.

3.4. Consistency of eye movement patterns across the three face tasks

The above results showed that face drawing expertise modulated eye movement pattern and performance in face matching but not in face recognition or part-whole tasks. This result suggests that face matching may have different task requirements from face recognition and part-whole tasks, whereas face recognition and part-whole tasks may have similar task requirements. Previous research has shown that when participants perform different tasks on the same stimuli, their eye gaze locations and transitions reflect differences in task requirements (e.g., Tatler, Wade, Kwan, Findlay, & Velichkovsky, 2010; Kanan, Bseiso, Ray, Hsiao, & Cottrell, 2015). Accordingly, participants may show consistent eye movement patterns between the face recognition and part-whole tasks, and this phenomenon may not be observed between face matching and face recognition/part-whole tasks. More specifically, those whose eye movement patterns were more eyes-focused in face recognition may also had more eyes-focused eye movement patterns in the part-whole task.[8] Consistent with this speculation, the results of Pearson correlation analysis showed that participants' eyes-nose scales in the face recognition task and the part-whole task were significantly correlated, $r(74) = 0.48$, $p < .001$, 95% CI [0.27, 0.67] (Fig. 14; $BF_{10} = 1583.00$, very strong evidence favoring the alternative hypothesis; percentage bend correlation, $r(74) = 0.51$, $p < .001$, 95% CI [0.30, 0.68]). In contrast, eyes-nose scale in the face matching task was not correlated with that in either the face recognition, $r(75) = -0.01$, $p = .92$, 95% CI [−0.25, 0.22] ($BF_{01} = 3.89$, positive evidence favoring the null hypothesis), or the part-whole task, $r(77) = -0.11$, $p = .33$, 95% CI [−0.34, 0.12] ($BF_{01} = 2.56$, weak evidence favoring the null hypothesis). This result suggests that face recognition and part-whole tasks may have similar task requirements that induced similar eye movement patterns, which differ from those of face matching.

In contrast, when we examined the consistency of eye movement patterns between the study and the test phases of the face recognition task, the eyes-nose scale during the study phase was significantly correlated with that during the test phase, $r(75) = 0.91$, $p < .001$, 95% CI [0.81, 1.0] (Fig. 14B; $BF_{10} = 3.25 \times 10^{25}$, very strong evidence favoring the alternative hypothesis; percentage bend correlation, $r(75) = 0.92$, $p < .001$, 95% CI [0.86, 0.95]). Similarly, in the part-whole task, the eyes-nose scale during the study phase was significantly correlated with that during the whole condition of the test phase, $r(77) = 0.82$, $p < .001$, 95% CI [0.69, 0.95] (Fig. 14C; $BF_{10} = 3.66 \times 10^{16}$, very strong evidence favoring the alternative hypothesis; percentage bend correlation, $r(75) = 0.79$, $p < .001$, 95% CI [0.64, 0.88]. Note that consistent with the test phase data, artists and novices did not differ significantly in eye movement pattern during the study phase of either the face recognition task, $t(73) = 0.64$, $p = .523$ (Equal variances assumed: Levene's test for equality of variances, $F(1,73) = 0.07$, $p = .796$. $BF_{01} = 2.67$, weak evidence favoring the null hypothesis), or the part-whole task, $t(76) = 0.44$, $p = .661$ (Equal variances assumed: Levene's test for equality of variances, $F(1,76) = 0.54$, $p = .464$. $BF_{01} = 3.00$, positive evidence favoring the null hypothesis).

4. Discussion

Here we investigated whether the effects of drawing expertise on face processing were task general and the possible underlying mechanisms through eye tracking and EMHMM analysis. Specifically, we examined the mechanism underlying portrait artists' limited advantage in face recognition regardless of their enhanced selective attention to the eyes/facial parts in addition to global face processing ability. Since portrait drawing involves discerning and matching identity-invariant, global and local facial information between two faces (the model and the drawing), an ability that is also required in a face matching task, we hypothesized that their advantage in face processing may be observed in simultaneous face matching but not in face recognition tasks, and this effect may be reflected in their eye gaze locations and transition behavior. More specifically, portrait artists may have better performance and a more eyes-focused eye movement pattern in a face matching task but not in face recognition tasks.

4.1. Effects of drawing expertise on face processing

Consistent with the hypothesis, our results showed that portrait artists adopted a more eyes-focused eye movement pattern (with

[7] One participant's eye movement data was missing due to technical problem.

[8] Note that since participants' eye movement data in the three experiments were analyzed separately, the eyes-focused and nose-focused patterns discovered through clustering in the three tasks were not identical. Instead, they reflected eye movement patterns that were specific to the individual tasks.

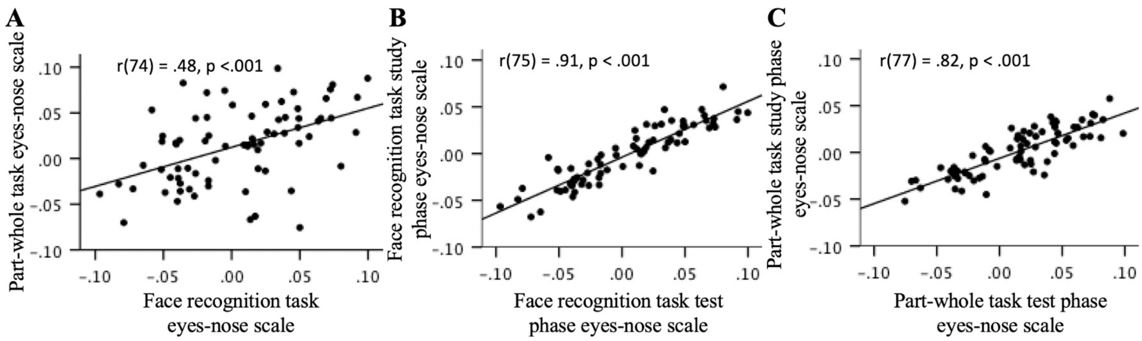


Fig. 14. (A) Participants' eye movement patterns (as measured in eyes-nose scale) in the face recognition task and the part-whole task were significantly correlated. Also, participants' eye movement patterns during the study and the test phase of the (B) face recognition task, and the (C) part-whole task, were significantly correlated.

positive evidence) and outperformed novices in simultaneous face matching (with very strong evidence), and participants' face drawing rating was positively correlated with both face matching performance and eye movement pattern similarity to the eyes-focused pattern (with positive evidence). This result suggested that portrait artists' face matching was associated with more selective attention to the eyes/facial parts than novices, and that this phenomenon and their better face matching performance were related to their drawing expertise. In the literature, portrait artists' advantage in the accuracy of perceptual discrimination between two computer-generated facial stimuli (Devue & Barsics, 2016) and reduced holistic face processing as assessed using the composite paradigm (suggesting better selective attention to facial parts; Zhou, Cheng, Zhang, & Wong, 2012) have been reported. Consistent with these previous findings, here we further showed that portrait artists looked at the eye region more often in perceptual discrimination between two faces. Note however that in Devue and Barsics (2016), as compared with novices, portrait artists had higher accuracy but longer response times than novices, suggesting a speed-accuracy tradeoff (a similar phenomenon was also observed in our face recognition task). In contrast, in the current study, artists had higher discrimination d' than novices but did not differ from novices in response times, suggesting better discrimination ability. This phenomenon may be because real face stimuli that preserved natural featural and configural variations among faces were used in the current study, which matched better with artists' face drawing experience, in contrast to computerized stimuli used in Devue and Barsics (2016).

In contrast, in face recognition, eye movement patterns of portrait artists did not differ significantly from novices (with positive evidence). This finding suggested that they did not engaged more selective attention to facial parts than novices, in contrast to the findings in the face matching task. This result was consistent with the hypothesis that their increased selective attention to facial parts as compared with novices may be limited to face processing tasks that were similar to their drawing experience such as simultaneous face matching. In addition, in recognition performance, we observed a speed-accuracy tradeoff in the participants. More specifically, as compared with novices, portrait artists had higher recognition sensitivity as measured in d' at the expense of longer response times and more eye fixations (with weak to positive evidence). A similar speed-accuracy tradeoff was also observed in previous studies comparing portrait artists and novice in sequential and simultaneous matching of faces (Devue & Barsics, 2016; Zhou et al., 2012). Together with the findings from eye movement pattern analysis, these results suggested that in face recognition portrait artists might be more motivated to achieve high accuracy and thus spent more time and efforts viewing the faces. Nevertheless, they did not necessarily use better eye movement patterns or have better ability for retrieving diagnostic information for face recognition. This finding is consistent with previous research showing that although participants with face drawing experience/training typically had better perceptual discrimination abilities for faces, they did not outperform non-drawers in face recognition (e.g., Devue & Barsics, 2016; Tree et al., 2017). One possibility is the additional memory demands in the face recognition task as compared with face matching may compete for cognitive recourses with the demands of selective attention to local facial features. Consequently, portrait artists' advantage in selective attention is compromised due to high cognitive demands in the face recognition task. Alternatively, it may be that the diagnostic features required for face matching and face recognition differ significantly. While face matching only requires discrimination of corresponding face-identity-invariant information between two faces, recognizing a face requires retrieving face-identity-invariant information of the face that is also idiosyncratic and diagnostic for distinguishing it from all other face identities (Rossion, 2018). This ability typically relies on one's statistical knowledge of face representations developed through life-long experience. An artist who received face drawing training typically has abundant experience in discerning and copying identity-invariant facial features in detail in order to make a genuine portrait drawing (Miall & Tachalenko, 2001; Tchalenko et al., 2003), but this experience did not necessarily help them acquire the statistics of diagnostic features for distinguishing a face from all other faces. Thus, although recent studies have suggested a significant genetic contribution in accounting for individual differences in face recognition ability (Shakeshaft & Plomin, 2015; Wilmer et al., 2010; Zhu et al., 2010), and face drawing experience does not seem to significantly enhance face recognition ability (Devue & Barsics, 2016; Dolzycka et al., 2014; Tree et al., 2017), it remains possible that a long-term experience in identifying diagnostic features among unfamiliar faces may enhance face recognition ability. One possibility may be professional caricaturists, who have learned to exaggerate distinctive characteristics of a face that can potentially be used for face recognition and identification. Future work will examine this possibility.

Similarly, in the part-whole task, a task that also required participants to remember face stimuli for a later recall, portrait artists and novices did not differ in the advantage of recognizing a whole face over recognizing a face part (i.e., the whole-part difference; with weak evidence) or in eye movement pattern when viewing whole-face stimuli (with positive evidence), although the eyes-focused pattern was associated with longer RT in the whole condition relative to the part condition (with positive evidence). This result again suggested that portrait artists' enhanced selective attention to face parts from their drawing experience was limited to face processing tasks similar to their drawing experience such as face matching; it did not generalize to face recognition memory tasks. Consistent with this speculation, while participants' eye movement patterns in the face recognition and part whole tasks were significantly correlated with each other (with very strong evidence), suggesting similarity in task requirements on face memory, eye movement pattern during face matching did not correlate with either that in face recognition or the part-whole task (with weak to positive evidence). Note that in the part-whole task, consistent with the findings in the face recognition task, portrait artists had longer response times and made more eye fixations to the faces than novices (with positive evidence), suggesting that they may be more motivated to perform the task than novices. In the literature, the part-whole effect has been considered as reflecting one aspect of holistic face processing that did not correlate well with the holistic processing effect measured using the composite paradigm (Rezlescu et al., 2017; Richler, Palmeri, & Gauthier, 2012). Indeed, whereas Zhou et al. (2012) showed that art students had reduced holistic face processing as compared with novices using the composite paradigm, here we did not observe a similar reduction in holistic face processing using the part-whole paradigm. Differences in task requirements may account for this inconsistency. In the composite task, participants judge whether the top halves of two presented faces are the same or different. Similar to a face matching task, it requires discrimination of corresponding facial part information between two faces, and thus may benefit from face drawing experience. In contrast, although the part-whole task also involves judgments between two faces, it requires retrieving diagnostic features for distinguishing the target face from other faces instead of simply comparing two presented faces as in the composite task. This ability may not be acquired naturally from face drawing experience. Thus, portrait artists' reduced holistic face processing as compared with novices may only be observed using the composite task, but not the part-whole paradigm.

Note however that although the artist and control groups recruited in the current study were matched as closely as possible, it remained possible that they could differ in other factors in addition to drawing experience. For example, they may differ in frequency of exposure to portraits or other types of art, or other characteristics relevant to their decision to receive drawing training or not. A longitudinal study that tracks the development of drawing skills is required to rule out these possibilities (Chamberlain et al., 2019).

4.2. Relations between gaze patterns and performance

Interestingly, in the face recognition task, we observed an inverted-U shaped, quadratic relationship between eye movement pattern and sensitivity performance measured in d', where the best performance was observed for eye movement patterns in between the eyes-focused and nose-focused patterns. Previous face recognition studies using the EMHMM approach have identified two representative eye movement patterns: an eyes-nose pattern in which eye fixations landed at both the eyes and the nose region, and a nose-focused pattern where eye fixations mainly centered at the nose region (Chuk et al., 2017; Chuk, Crookes, et al., 2017, Chan et al., 2018). These studies have consistently shown a linear relationship between eye movement pattern and recognition performance: the higher the similarity to the eyes-nose pattern, the better the performance. Miellet et al. (2011) reported that in Caucasian participants, looking at the eyes during face perception was associated with local/high spatial frequency information processing whereas looking at the nose associated with global/low spatial frequency information processing (relative to the face size). Consistent with this finding, Cheng et al. (2015) found that in Asian participants, local attention priming using hierarchical letter stimuli made participants' eye movement patterns more similar to the eyes-nose pattern as opposed to the nose-focused pattern and increased their recognition performance, as compared with no priming or global attention priming conditions (Cheng et al., 2015). These findings demonstrated the association between engagement of local/global attention and the eyes-nose/nose-focused eye movement pattern across culture. Indeed, Chuk, Crookes, et al. (2017) showed that the advantage of the eyes-nose pattern was also observed across cultures. This result suggested that the flexibility to engage both local and global facial information processing may be optimal for face recognition (Chuk et al., 2017; Chuk, Crookes, et al., 2017). Indeed, in the literature, although holistic face processing has been suggested to be a result of our experience and expertise in face recognition, whether holistic processing predicts face recognition performance has not been consistently reported (e.g., Konar, Bennett, & Sekuler, 2010; Rezlescu et al., 2017; Richler, Cheung, & Gauthier, 2011; Richler, Floyd, & Gauthier, 2015; Verhallen et al., 2017; Wang, Li, Fang, Tian, & Liu, 2012). This inconsistency may be because the holistic process effect assessed in the composite paradigm reflects failure of selective attention to local features, whereas optimal face recognition performance requires both global and local face processing skills.

Note that the eyes-focused eye movement pattern discovered in the current study differed from the eyes-nose pattern observed in the previous studies (Chuk et al., 2017; Chuk, Crookes, et al., 2017, Chan et al., 2018) in that it focused mainly at the eye region (whereas the nose-focused pattern corresponded well to those observed in the previous studies), allowing the quadratic relationship between eye movement pattern and face recognition performance to be observed. Since the discovery of representative eye movement patterns from the participants through the clustering method in EMHMM is entirely data driven,[9] this difference between the current study and the previous studies may be due to difference in participant sample or stimulus used. In contrast to other eye movement data analysis methods such as using predefined regions of interests (ROIs) or fixation heat maps, EMHMM derives personalized ROIs and transition probabilities among the ROIs for each individual directly from data. Thus, it accounts for individual differences in both spatial and temporal dimension of eye movements (i.e., eye gaze locations and transitions) better than previous approaches. More importantly, it provides quantitative measures of eye movement pattern similarities (incorporating both spatial and temporal dimensions) among individuals through machine learning methods (i.e., the log-likelihood measures), enabling us to discover the relationship between eye movement pattern and face recognition performance.

In contrast to face recognition, in face matching although higher drawing rating was associated with both better face matching performance and more eyes-focused eye movement pattern, no significant correlation was observed between eye movement pattern and face matching performance (with weak evidence). This result suggested that portrait artists' advantage in face matching could not be simply explained by their more eyes-focused eye movement patterns. As shown in Fig. 6A and Fig. 7, although portrait artists had better face matching performance and more eyes-focused eye movement patterns than novices, there was a larger variance in eye movement pattern in novices than artists. As the result, some novices adopted a similar level of eyes-focused pattern to artists but had lower face matching performance than artists. This phenomenon suggested that in addition to engaging a more eyes-focused eye movement pattern, portrait artists may have better ability to extract face-identity-invariant information important for discriminating between two faces. Although some novices also engaged eyes-focused eye movement patterns, they performed worse than artists because of lack of this ability. Consistent with this speculation, participants' performance in face matching was significantly correlated with their drawing rating, which may have reflected their ability to extract face-identity-invariant information between two faces during drawing or face matching. In addition, when we examined portrait artists' and novices' data separately, this correlation was significant in portrait artists (with positive evidence) but marginal in novices (with weak evidence), suggesting that it may be related to portrait drawing experience.

4.3. Comparisons with other expertise domains

In contrast to face recognition, writing practice has been consistently reported to enhance Chinese character recognition performance (e.g., Tso et al., 2014; McBride-Chan et al., 2011; Tan et al., 2005). Note

[9] That is, the clustering is based on only the eye fixation location and transition information obtained from eye tracking.

however that Chinese character recognition involves only the recognition of a fixed number of characters (about 3000 to 4000 for a proficient reader) that are different combinations of a limited number of stroke patterns/components (about 200; Hsiao & Shillcock, 2006). A Chinese reader typically has to learn to read and write each character one by one, and Chinese character recognition performance is typically measured using existing characters with components that are already familiar to the readers. Thus, Chinese character recognition tasks used in the literature are analogous to familiar face identification instead of unfamiliar face recognition. Familiar face recognition is shown to be more accurate and automatized than unfamiliar face recognition, as familiarity helps us distinguish within-person variability from between-person variability, whereas in unfamiliar face recognition these two types of variability are easily confusable (Young & Burton, 2018). Since face drawing involves discerning and copying identity-invariant information from a face, it may facilitate familiarization of the face to make its identification more accurate and automatized. Writing practice may have a similar effect on Chinese character identification, consistent with the literature. It remains unclear whether writing experience with Chinese characters facilitates recognition of unfamiliar characters. Indeed, Liu, Chuk, Yeh, and Hsiao (2016) showed that the reduced holistic processing effect associated with expert Chinese character recognition did not generalize well across simplified and traditional Chinese characters. A similar phenomenon was observed in object recognition, where car experts' holistic processing in the perception of modern cars did not generalize to antique cars (Bukach, Phillips, & Gauthier, 2010). These results suggested that writing experience with Chinese characters, which enhances readers' processing of local features of characters (Tso et al., 2014), does not necessarily generalize to or facilitate the perception of unfamiliar characters, consistent with the current finding about the effect of drawing experience on (unfamiliar) face recognition. Future work will examine this possibility.

In conclusion, through eye movement analysis with EMHMM, here we show that face drawing experience is associated with increased selective attention to the eyes as reflected in eye movement pattern. However, this effect is limited to tasks similar to their drawing experience such as face matching, and does not generalize to face recognition tasks. This result may be due to different diagnostic information required between face matching and face recognition tasks: face matching requires abilities for extracting and comparing identity-invariant facial features between two faces, which may be naturally acquired through face drawing practice in order to make a genuine portrait. In contrast, face recognition depends on identity-invariant information that is also distinct from other faces, which may require more specialized training than simple face drawing. We also show that for face recognition, eye movement patterns that involve a mixture of eyes- and nose-focused patterns are optimal for performance, suggesting that optimal face processing performance typically involves a combination of local and global facial information processing. In contrast, participants' performance in face matching was predicted by participants' drawing rating but not eye movement pattern, suggesting that artists' advantage in face matching may be related to their better ability in extracting identity-invariant facial features between two faces developed through drawing experience rather than more eyes-focused eye movement patterns. Thus, although recent research has shown a substantial genetic contribution in face recognition ability, it remains possible that long-term experience with discerning and extracting identity-invariant and person-specific information that are diagnostic for unfamiliar face recognition may enhance face recognition ability.

Acknowledgment

We are grateful to the Research Grant Council of Hong Kong (project code: #17402814 and #17609117 to J. H. Hsiao). The work is also supported by a Strategic Research Grant from City University of Hong Kong (Project No. 7005218 to A. B. Chan). We thank Sara T. K. Li and Hui Fei Chan for their help on an early version of the study. We thank the Editor and three anonymous reviewers for the helpful comments.

References

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology, 25*(3), 281–316.
Barton, J. J., Keenan, J. P., & Bass, T. (2001). Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration. *British Journal of Psychology, 92*(3), 527–549.
Birmingham, E., Bischof, W. F., & Kingstone, A. (2008a). Social attention and real-world scenes: The roles of action, competition and social content. *The Quarterly Journal of Experimental Psychology, 61*(7), 986–998.
Birmingham, E., Bischof, W. F., & Kingstone, A. (2008b). Gaze selection in complex social scenes. *Visual Cognition, 16*(2–3), 341–355.
Bonnen, K., Klare, B., & Jain, A. (2013). Component-based representation in automated face recognition. *IEEE Transactions on Information Forensics and Security, 8*(1), 239–253.
Bukach, C. M., Phillips, S. W., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Performance and Psychophysics, 72*(7), 1865–1874.
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. *Psychological Science, 11*(5), 429–433.
Caldara, R., & Miellet, S. (2011). iMap: A novel method for statistical fixation mapping of eye movement data. *Behavior Research Methods, 43*(3), 864–878.
Chamberlain, R., Drake, J. E., Kozbelt, A., Hickman, R., Siev, J., & Wagemans, J. (2019). Artists as experts in visual cognition: An update. *Psychology of Aesthetics, Creativity, and the Arts, 13*(1), 58–73.
Chamberlain, R., McManus, I. C., Riley, H., Rankin, Q., & Brunswick, N. (2013). Local processing enhancements associated with superior observational drawing are due to enhanced perceptual functioning, not weak central coher- ence. *The Quarterly Journal of Experimental Psychology, 66*, 1448–1466.
Chamberlain, R., & Wagemans, J. (2015). Visual arts training is linked to flexible attention to local and global levels of visual stimuli. *Acta Psychologica, 161*, 185–197.
Chan, C. Y., Chan, A. B., Lee, T. M., & Hsiao, J. H. (2018). Eye-movement patterns in face recognition are associated with cognitive decline in older adults. *Psychonomic Bulletin & Review*, 1–8.
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*, 154.
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.
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.
Coviello, E., Chan, A. B., & Lanckriet, G. R. G. (2014). Clustering hidden Markov models with variational HEM. *Journal of Machine Learning Research, 15*(Feb), 697–747.
Davis, J., McKone, E., Zirnsak, M., Moore, T., O'Kearney, R., Apthorp, D., & Palermo, R. (2017). Social and attention-to-detail subclusters of autistic traits differentially predict looking at eyes and face identity recognition ability. *British Journal of Psychology, 108*(1), 191–219.
Devue, C., & Barsics, C. (2016). Outlining face processing skills of portrait artists: Perceptual experience with faces predicts performance. *Vision Research, 127*, 92–103.
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* (pp. 280–284).
Dolzycka, D., Herzmann, G., Sommer, W., & Wilhelm, O. (2014). Can training enhance face cognition abilities in middle-aged adults? *PLoS One, 9*(3), Article e90249.
Drewes, J., Zhu, W., Hu, Y., & Hu, X. (2014). Smaller is better: Drift in gaze measurements due to pupil dynamics. *PLoS One, 9*(10), Article e111197.
Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482–498.
Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29*(2), 159–170.
Gao, Y., Uozumi, T., & Chen, F. (2005). Off-line handwritten Chinese character recognition based on fusion features and Bayesian classifier. In *Soft Computing as Transdisciplinary Science and Technology* (pp. 563–570). Springer.
Galmar, B., Chung, H., Hsiao, J. H., Chuk, T., Chan, A. B., & Hsiao, J. H. (2014). Face drawing experience is associated with better face recognition performance and reduced left-side bias in face perception. *Journal of Vision, 14(10)*, 1260.
Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41*(17), 2261–2271.
Guan, C. Q., Harris, L. N., Meng, W., & Perfetti, C. A. (2015). Writing quality predicts Chinese learning. *Reading and Writing: An Interdisciplinary Journal, 28*(6), 763–795.
Harris, A., & Aguirre, G. K. (2008). The representation of parts and wholes in face-selective cortex. *Journal of Cognitive Neuroscience, 20*(5), 863–878.
Harris, A., & Nakayama, K. (2008). Rapid adaptation of the M170 response: Importance of face parts. *Cerebral Cortes, 18*(2), 467–476.

Hayward, W. G., Crookes, K., & Rhodes, G. (2013). The other-race effect: Holistic coding differences and beyond. *Visual Cognition, 21*, 1224–1247.
Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33*(1), 98–106.
Hole, G. J. (1994). Configurations! Factors in the perception of unfamiliar faces. *Perception, 23*(1), 65–74.
Hooge, I. T., Hessels, R. S., & Nyström, M. (2019). Do pupil-based binocular video eye trackers reliably measure vergence? *Vision Research, 156*, 1–9.
Hsiao, J. H., Cottrell, G. W., Chuk, T., Chan, A. B., & Hsiao, J. H. (2009). Not all expertise is holistic, but it may be leftist: The case of Chinese character recognition. *Psychological Science, 20(4)*, 455–463.
Hsiao, J. H., & Shillcock, R. (2006). Analysis of a Chinese phonetic compound database: Implications for orthographic processing. *Journal of Psycholinguistic Research, 35*, 405–426.
Jain, A. K., Chen, Y., & Demirkus, M. (2007). Pores and ridges: High-resolution fingerprint matching using level 3 features. *IEEE Transactions on Pattern Analysis, 29*(1), 15–27.
Joseph, R. M., & Tanaka, J. (2002). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry, 43*(8), 1–14.
Kanan, C., Bseiso, D., Ray, N., Hsiao, J. H., & Cottrell, G. (2015). Humans have idiosyncratic and task-specific scanpaths for judging faces. *Vision Research, 108*, 67–76.
Kemp, R., McManus, C., & Pigott, T. (1990). Sensitivity to the displacement of facial features in negative and inverted images. *Perception, 19*(4), 531–543.
Konar, Y., Bennett, P. J., & Sekuler, A. B. (2010). Holistic processing is not correlated with face-identification accuracy. *Psychological Science, 21*, 38–43.
Kozbelt, A., & Ostrofsky, J. (2018). Expertise in drawing. In H. Ericsson, & W. Kozbelt (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 576–596). Cambridge University Press.
Lan, H., Liu, Z., Hsiao, J. H., Yu, D., & Chan, A. B. (submitted). Clustering Hidden Markov Models with Variational Bayesian Hierarchical EM.
Lau, E. Y. Y., Ip, M. S. M., Lee, T. M. C., Yeung, A. W. M., & Eskes, G. A. (2016). Neurocognitive and psychosocial outcomes of obstructive sleep apnea in Hong Kong Chinese: Similar to or different from western populations? *Journal of Sleep Disorders: Treatment and Care, 2*, 3.
Lee, S. H., & Kim, J. H. (2008). Complementary combination of holistic and component analysis for recognition of low-resolution video character images. *Pattern Recognition Letters, 29*(4), 383–391.
Liang, F., Paulo, R., Molina, G., Clyde, M. A., & Berger, J. O. (2008). Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association, 103*(481), 410–423.
Liu, T., Chuk, T., Yeh, S. L., & Hsiao, J. H. (2016). Transfer of perceptual expertise: The case of simplified and traditional Chinese character recognition. *Cognitive Science, 40*(8), 1941–1968.
Lobmaier, J. S., & Mast, F. W. (2007). Perception of novel faces: The parts have it! *Perception, 36*(11), 1660–1673.
Longcamp, M., Anton, J. L., Roth, M., & Velay, J. L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *Neuroimage, 19*, 1492–1500.
Ma, L.-L., & Wu, J. (2012). A Component-based On-line Handwritten Tibetan Character Recognition Method Using Conditional Random Field. In *Paper presented at the International Conference on Frontiers in Handwriting Recognition (ICFHR)*.
McBride-Chang, C., Chung, K. K. H., & Tong, X. (2011). Copying skills in relation to word reading and writing in Chinese children with andwithout dyslexia. *Journal of Experimental Child Psychology, 110*, 422–433.
Mehoudar, E., Arizpe, J., Baker, C., & 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.
Miall, R. C., & Tchalenko, J. (2001). A painter's eye movements: A study of eye and hand movement during portrait drawing. *Leonardo, 34*(1), 35–40.
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.
Mohr, B., Pulvermüller, F., & Zaidel, E. (1994). Lexical decision after left, right and bilateral presentation of function words, content words and non-words: Evidence for interhemispheric interaction. *Neuropsychologia, 32*(1), 105–124.
Mondloch, C. J., Elms, N., Maurer, D., Rhodes, G., Hayward, W. G., Tanaka, J. W., & Zhou, G. (2010). Processes underlying the cross-race effect: An investigation of holistic, featural, and relational processing of own-race versus other-race faces. *Perception, 39*, 1065–1085.
Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353–383.
O'Toole, A. J., Deffenbacher, K. A., Valentin, D., & Abdi, H. (1994). Structural aspects of face recognition and the other-race effect. *Memory and Cognition, 22*, 208–224.
Ostrofsky, J., Cohen, D. J., & Kozbelt, A. (2014). Objective versus subjective measures of face-drawing accuracy and their relations with perceptual constancies. *Psychology of Aesthetics Creativity and the Arts, 8(4)*, 486.
Ostrofsky, J., Kozbelt, A., Cohen, D. J., Conklin, L., & Thomson, K. (2016). Face inversion impairs the ability to draw long-range, but not short-range, spatial relationships between features. *Empirical Studies of the Arts, 34(2)*, 221–233.
Ostrofsky, J., Pletcher, R., & Smith, J. (2020). The Effects of Disrupting Holistic Processing on the Ability to Draw a Face. *Art & Perception, 8(1)*, 68–88.
Perdreau, F., & Cavanagh, P. (2013). The artist's advantage: Better integration of object information across eye move- ments. *i-Perception, 4*, 380–395.
Perdreau, F., & Cavanagh, P. (2015). Drawing experts have better visual memory while drawing. *Journal of Vision, 15*(5), Article 5.
Pernet, C. R., Wilcox, 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. (2012). Optimal eye movements for face recognition. *Proceedings of the National Academy of Sciences, 109*(48), E3314–E3323.
Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 111–163.
Rezlescu, C., Susilo, T., Wilmer, J. B., & Caramazza, A. (2017). The inversion, part-whole, and composite effects reflect distinct perceptual mechanisms with varied relationships to face recognition. *Journal of Experimental Psychology: Human Perception and Performance, 43*(12), 1961–1973.
Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science, 22*, 464–471.
Richler, J. J., Floyd, R. J., & Gauthier, I. (2015). About-face on face recognition ability and holistic processing. *Journal of Vision, 15*(9), 1–12.
Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology, 3*, 553.
Rossion, B. (2018). Humans are visual experts at unfamiliar face recognition. *Trends in Cognitive Sciences, 22*(6), 471–472.
Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225–237.
Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial–relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance, 22*(4), 904.
Shakeshaft, N. G., & Plomin, R. (2015). Genetic specificity of face recognition. In *, 112. PNAS Proceedings of the National Academy of Sciences of the United States of America* (pp. 12887–12892).
Siok, W. T., Perfetti, C. A., Jin, Z., & Tan, L. H. (2004). Biological abnormality of impaired reading is constrained by culture. *Nature, 431*, 71–76.
Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *PNAS, 102*, 8781–8785.
Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology, 46*(2), 225–245.
Tatler, B., Wade, N., Kwan, H., Findlay, J., & Velichkovsky, B. (2010). Yarbus, eye movements, and vision. *i-Perception, 1*, 7–27.
Tchalenko, J., Dempere-Marco, L., Hu, X. P., & Yang, G. Z. (2003). Eye movement and voluntary control in portrait drawing. In J. Hyona, R. Radach, & H. Deubel (Eds.), *The Mind's eye: Cognitive and applied aspects of eye movement research* (pp. 705–727). Amsterdam: Elsevier.
Tree, J. J., Horry, R., Riley, H., & Wilmer, J. B. (2017). Are portrait artists superior face recognizers? Limited impact of adult experience on face recognition ability. *Journal of Experimental Psychology: Human Perception and Performance, 43*(4), 667.
Tso, R. V. Y., Au, T. K., & Hsiao, J. H. (2012). Writing facilitates learning to read in Chinese through reduction of holistic processing: A developmental study. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual conference of the cognitive science society* (pp. 2463–2468). Austin, TX: Cognitive Science Society.
Tso, R. V. Y., Au, T. K., & Hsiao, J. H. (2014). Perceptual expertise: Can sensorimotor experience change holistic processing and left side bias? *Psychological Science, 25*(9), 1757–1767.
Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Bargary, G., & Mollon, J. D. (2017). General and specific factors in the processing of faces. *Vision Research, 141*, 217–227.
Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science, 23*, 169–177.
White, D., Phillips, P. J., Hahn, C. A., Hill, M., & O'Toole, A. J. (2015). Perceptual expertise in forensic facial image comparison. *Proceedings of the Biological Sciences, 282*(1814), Article 20151292.
Wilmer, J. B., Germine, L., Loken, E., Guo, X. M., Chatterjee, G., Nakayama, K., … Duchaine, B. (2010). Response to Thomas: Is human face recognition entirely genetic? *Proceedings of the National Academy of Sciences, USA, 107*. letter 24.
Young, A. W., & Burton, M. (2018). Are we face experts? *Trends in Cognitive Sciences, 22*(2), 100–110.
Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face per- caption. *Perception, 16*(6), 747–759.
Zhang, D., Zuo, W., & Yue, F. (2012). A comparative study of palmprint recognition algorithms. *ACM Computing Surveys, 44*(1), 2.
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*(2).
Zhang, L., Zhang, L., Zhang, D., & Zhu, H. (2011). Ensemble of local and global information for finger–knuckle-print recognition. *Pattern Recognition, 44*(9), 1990–1998.
Zhou, G., Cheng, Z., Zhang, X., & Wong, A. C. N. (2012). Smaller holistic processing of faces associated with face drawing experience. *Psychonomic Bulletin & Review, 19*(2), 157–162.
Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., … Liu, J. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology, 20*, 137–142.